Elektronica en Telecommunicatie 1

Gemeenschappelijke publikatie van de Sectie voor Telecommunicatietechniek van het K.I.v.I. en het Nederlands Elektronica- en Radiogenootschap.

Redactie-adres: Prinsessegracht 23, Den Haag.

'Liber amicorum',

opgedragen aan prof. dr. ir. J. P. Schouten



Prof. dr. ir. J. P. Schouten

Het afscheid op 30 januari 1970 van prof. dr. ir. J. P. Schouten als gewoon hoogleraar in de Afdeling der Elektrotechniek van de Technische Hogeschool te Delft, is voor velen een aanleiding om te getuigen van hun gevoelens van hoogachting en waardering. Als uiting hiervan hebben collega's en oud-leerlingen een bijdrage geleverd, die zij bij dezen aan hem opdragen. [1]

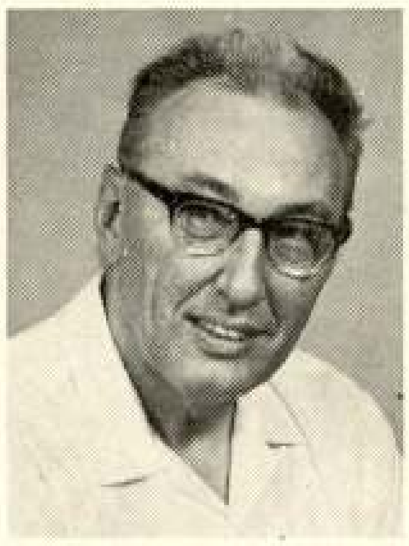
Dank zij de welwillende medewerking van de redactie van 'De Ingenieur' is het mogelijk deze artikelen te publiceren in deze aflevering van 'Elektronica en Telecommunicatie'.

Namens de commissie tot regeling van
het afscheid van prof. dr. ir. J. P. Schouten,

De Voorzitter,

Prof. ir. J. W. Alexander.

[1]) The following articles are dedicated to Professor J. P. Schouten on the occasion of his retirement.



517.52:621.372.57

Determination of the characteristic of a non-linear resistor by harmonic excitation

by prof. dr. C. J. Bouwkamp, Philips Research Laboratories,
N.V. Philips' Gloeilampenfabrieken, Eindhoven

Consider a non-linear resistor with centre-symmetric current-voltage characteristic $i = f(v) = -f(-v)$, and let this resistor be excited by a harmonic signal $v = x\cos(\omega t)$, where the amplitude x covers a certain range ($0 \leqq x \leqq a$). Assume further that the amplitude $g(x)$ of the first harmonic $g(x)\cos(\omega t)$ in the output current $f(x\cos(\omega t))$ is measured in the same range. Then the question is:

– Can the function $f(x)$ be determined uniquely from the function $g(x)$, where $0 \leqq x \leqq a$; and if so, how can it be done? The answer to this question is of importance in relation to certain measuring techniques in experimental physics [1] and is reported in this note.

If the resistor is excited by $v = x\cos\varphi$, where $\varphi = \omega t$, the output is $f(x\cos\varphi)$, and the amount of the first harmonic in this output clearly is

$$\frac{1}{\pi}\int_0^{2\pi} f(x\cos\varphi)\cos\varphi\,\mathrm{d}\varphi$$

In virtue of the symmetry of the cosine function and $f(v)$ being an odd function of v, we can replace the upper limit by $\pi/2$ and multiply the expression by 4. Therefore, the two functions f and g are interconnected by the following integral equation:

$$\int_0^{\pi/2} f(x\cos\varphi)\cos\varphi\,\mathrm{d}\varphi = \frac{\pi}{4}\,g(x), \qquad 0 \leqq x \leqq a$$

This integral equation is closely related to Abel's integral equation, as can be seen by introducing the new variable of integration $y = x\cos\varphi$.

On substitution we get

$$\int_0^x \frac{f(y)\,y\,\mathrm{d}y}{\sqrt{x^2-y^2}} = \frac{\pi}{4}\,x\,g(x)$$

The solution of this latter equation is known. Under certain mild conditions as to smoothness of f and g, not to be specified here, the integral equation has the unique solution

$$f\,(y) = \frac{1}{2y}\,\frac{\mathrm{d}}{\mathrm{d}y}\int_0^y \frac{g\,(x)\,x^2\,\mathrm{d}x}{\sqrt{y^2-x^2}} \qquad (0 \leq y \leq a)$$

and this answers our question.

This solution in closed form though is not very attractive to the practical engineer if g is not known analytically. An alternative solution by means of power series or the like is more practical, especially so if the resistor is slightly non-linear. Therefore, let us assume that

$$g\,(x) = \sum_n c_n\,x^{2n+1}$$

where the right-hand expression may be an infinite convergent series, or a least-square polynomial approximation to measured values of g, or a Chebyshev approximation, etc. If this is substituted in the closed-form solution, the result is an expression of the same type, namely

$$f\,(y) = \sum_n 2^{2n}\cdot\frac{n!\,(n+1)\,!}{(2n+1)\,!}\cdot c_n\,y^{2n+1}$$

where the summation is extended over the same set of values n as in the expression assumed for g. In the case of infinite series, the two power series have the same radius of convergence.

More explicitly, if the measured values of g can be represented by a polynomial approximation, say

$$g(x) \approx c_0 x + c_1 x^3 + c_2 x^5 + c_3 x^7$$

then the required characteristic is of the form

$$f(v) = c_0 v + \tfrac{4}{3} c_1 v^3 + \tfrac{8}{5} c_2 v^5 + \tfrac{64}{35} c_3 v^7$$

of course approximately, and over the same range.

Finally, it should be noted that in practice we need not invoke the theory of Abel's integral equation at all, apart perhaps from the question of uniqueness. In fact, if we take $f(v) = v^m$ $(v \geqq 0)$, where m is not necessarily an integer, then the corresponding $g(x)$, $x \geqq 0$, is apparently given by

$$x^{-m}\,g\,(x) = \frac{4}{\pi}\int_0^{\pi/2} \cos^{m+1}\varphi\;\mathrm{d}\varphi = \frac{\Gamma\left(1+\frac{m}{2}\right)}{\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{m}{2}+\frac{3}{2}\right)}$$

and therefore, if

$$g\,(x) \approx \sum_m b_m\,x^m \qquad (x \geqq 0,\ m > 0)$$

then

$$f\,(v) \approx \sum_m \frac{\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{3}{2}+\frac{m}{2}\right)}{\Gamma\left(1+\frac{m}{2}\right)}\,b_m\,v^m \qquad (v \geqq 0)$$

The author is indebted to Professor Butterweck, Technological University Eindhoven, for drawing his attention to this problem.

Reference

[1] G. A. Ackert and J. de Groot: Measurements of the current-field strength characteristic of n-type gallium arsenide using various high-power microwave techniques. Trans. I.E.E.E., **ED-14**, 505-511, 1967.

621.375

On the stability of a circuit with negative resistance with non-linear characteristic

by prof. ir. J. W. Alexander, Delft University of Technology, Dept. of Electronical Engineering



Summary: For an amplifier with a negative resistance the operating point on the curved characteristic of the element has to be stable. A criterion for the stable points on the curved characteristic is given, without the addition of an extra capacitance or inductance. This criterion runs as follows: $\mathrm{d}E/\mathrm{d}I$ must be positive, where E and I are the adjustment voltage and ditto current respectively. When the circuit contains also other elements than only resistances the amplifier circuit is stable when in addition to the adjustment stabilization the poles must have a negative real part.

1. Introduction

A negative resistance always has a curved characteristic with a negative slope on a part of the curve. For operation as an amplifier we have to choose our operating at any point on the negative slope of the characteristic. For stable operation this operating point has to be a stable one. In the discussion on this problem one usually takes into account the existence of an extra element, either an inductance or a capacitance. In the following an other method is used, without the addition of an extra element.

2. Criterion of stability

We will start from the fact that a circuit consisting of a d.c. source with voltage E and a positive resistance R is always

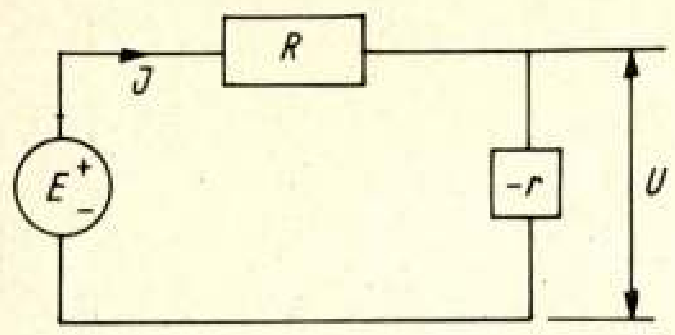


Fig. 1. Circuit with a positive and a linear negative resistance.

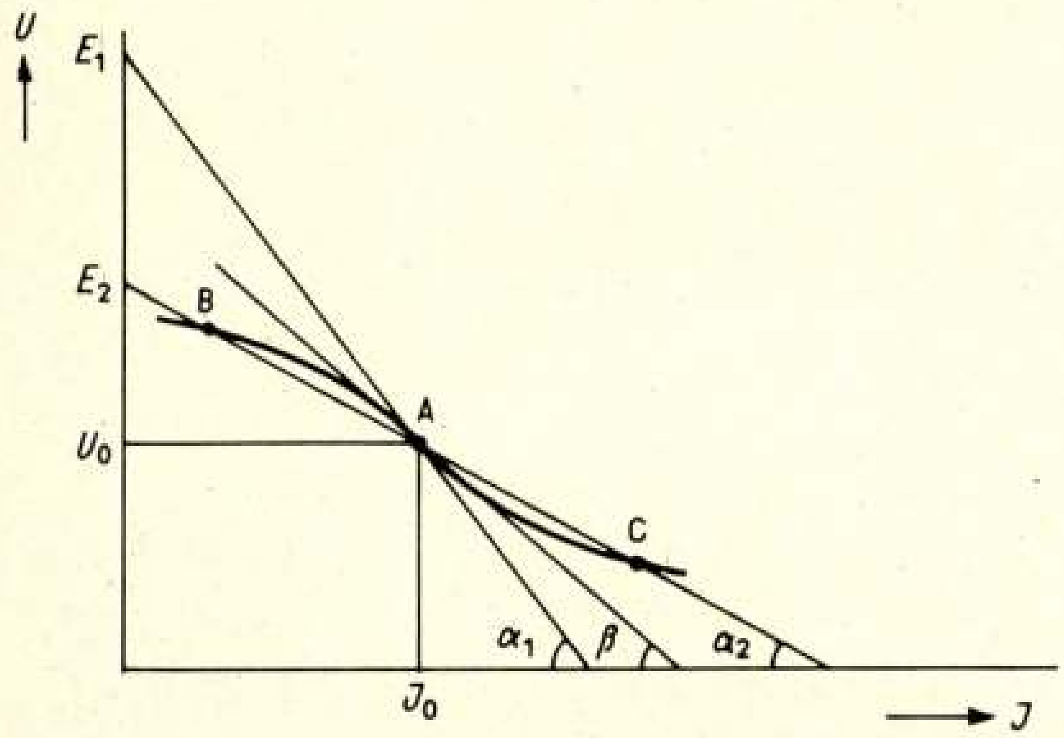


Fig. 2. Characteristic of a non-linear negative resistance with different load lines.

stable. That means that a small change dE in the voltage E gives a small change dI in the current, in such a way that $dE/dI = +R$, a positive value.

For a linear negative resistance $-R$ we should have $E = -IR$, so $dE/dI < 0$. This is a negation of the stable situation $dE/dI > 0$, consequently an unstable situation, in agreement with the impossible situation $E = -IR$.

When the circuit consists of a positive and a negative resistance in series, according to Fig. 1, then $dE/dI = R - r$.

When $R > r$, dE/dI is positive and the circuit is stable, when $R < r$ however, it is unstable.

3. Circuit with a non-linear negative resistance

A negative resistance always has a curved characteristic with a negative slope on a part of the curve. Let Fig. 2 represent the U vs. I-diagram of the non-linear element.

In this figure A is the operating point chosen in the inflexion point of the characteristic, where

$$E_1 - R_1 I_0 = U_0 \tag{1}$$

where E_1 is the adjustment voltage and $\tan(\alpha_1) = R_1$.

In the figure another situation is also shown, where

$$E_2 - R_2 I_0 = U_0 \text{ with } \tan(\alpha_2) = R_2$$

With

$$U = U_0 + u \text{ and } I = I_0 + i \tag{2}$$

we may describe the curved characteristic by the equation

$$u = -ri + bi^3 \tag{3}$$

from A till beyond the extreme values. Here r and b are positive and $\tan(\beta) = r$. We see in the figure that $R_1 > r$ and $R_2 < r$.

The points of intersection of the characteristic and the load line follow in the first situation from

$u = -ri + bi^3$ together with $u = -R_1 i$.

This results in $i = 0$ and $-R_1 = -r + bi^2$, or $bi^2 = r - R_1$. Now $R_1 > r$, so we have only one intersection point A.

For the second case we find in the same way

$i = 0$ and $bi^2 = r - R_2$

or the three intersection points A, B and C.

For the discussion of the stability we proceed as follows:
In the first situation Fig. 1 shows that

$$E_1 = R_1 I + U \tag{4}$$

Substituting (1), (2) and (3) in (4) we obtain

$$E_1 - I_0 R_1 - U_0 = (R_1 - r)\, i + bi^3.$$

For the adjustment stability we get

$$dE/dI = dE_1/di = R_1 - r + 3\,bi^2$$

for I_0 and U_0 are constant.

The only intersection point is $i = 0$. As $R_1 > r$ the point A is stable.

But for the second case with $R_2 < r$ we get

$$dE/dI = dE_2/di = R_2 - r + 3\,bi^2.$$

For A: $i = 0$ and $R_2 < r$, consequently A is unstable.

For B and C: $bi^2 = r - R_2$, so

$$dE/dI = R_2 - r + 3\,(r - R_2) = 2\,(r - R_2).$$

As $R_2 < r$, B and C are stable points.

Consequently: For operating in the inflection point A the positive resistance R must be greater than the negative resistance r.

4. Situation with a negative conductance

Till now we have studied the element with the negative resistance, according to Fig. 2, or the element with the current-controlled characteristic.

The other element with the voltage-controlled characteristic, i.e. the same form as in Fig. 2, but with U and I interchanged, has the criterion of adjustment stability: $dA/dU > 0$, where A is the adjustment current. This results in the similar condition that for operating in the inflection point A the positive conductance G must be greater than the negative conductance g of the element.

5. Amplification with a negative resistance and other elements

In this case the requirement for the adjustment stability is not the only one which has to be satisfied. The poles of the characteristic equation of the circuit must have a negative real part, or the Hurwitz criteria are to be satisfied also.

To demonstrate this we examine the circuit, shown in Fig. 3,

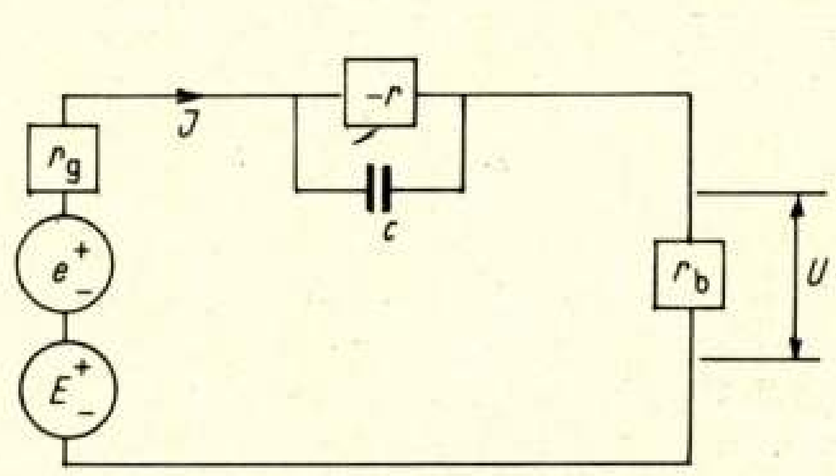


Fig. 3. Circuit for amplification with a current-controlled negative resistance.

for amplification with a current-controlled negative resistance.

For the adjustment stability we get from $I = E/(r_g + r_b - r)$ that

$$r_g + r_b > r \tag{5}$$

is required.

For the time-depending e we get from

$$u = e\, r_b\,(1 - rpC)/\{r_g + r_b - r - (r_g + r_b)\, rpC\}$$

the characteristic equation:

$r_g + r_b - r - (r_g + r_b)\, rpC = 0.$

For a stable situation it is well-known that the pole must be negative, or

$$r_g + r_b < r \tag{6}$$

Now (5) and (6) cannot be realized at the same time, consequently: the circuit according to Fig. 3 with a current-controlled negative resistance is not possible as an amplifier.

621.373.5

Niet-lineariteit van de lawine-looptijd-oscillator

door dr. ir. K. Mouthaan, Natuurkundig Laboratorium,
N.V. Philips' Gloeilampenfabrieken, Eindhoven



Summary: *Non-linearity of the Avalanche Transit-Time Oscillator.*
High-power solid-state generation of microwaves has become possible with the Avalanche Transit-Time Oscillator. The oscillator is based on the dynamic negative-resistance properties of semiconductor diodes that are reverse-biased into avalanche breakdown. Optimum design conditions are sought by analysis of the non-linear operating characteristics. Simultaneous operation at two harmonically related frequencies is shown to permit well-defined design of the oscillator, and to yield high output power and low frequency-modulation noise. Theoretical tuning conditions and curves for output power as a function of the bias current are presented.

De lawine-looptijd-oscillator is een recent tot ontwikkeling gekomen vaste-stof microgolfbron. Frequenties van 1 GHz tot meer dan 300 GHz en nuttige vermogens tot enige watts in continu bedrijf zijn experimenteel gerealiseerd.

Voorgesteld door Read [1], berust de oscillator op looptijdnegatieve weerstand van een in lawinedoorslag bedreven halfgeleiderdiode. Een succesvolle diodestructuur is die met asymmetrische abrupte p-n-overgang. Het profiel van het elektrische veld binnen een dergelijke in sperrichting aangesloten structuur is schematisch voorgesteld in fig. 1. Het verontreinigingsprofiel is zo dat bij voldoend hoge sperspanning enerzijds lawinedoorslag optreedt op de p^+-n-overgang ten gevolge van stootionisatie; anderzijds strekt de depletielaag zich dan uit tot aan de n^+-laag. In de gehele depletielaag is daarbij de veldsterkte voldoende hoog om de in de doorslag gegenereerde elektronen de n-laag met veldverzadigde driftsnelheid te doen doorlopen. De tevens gegenereerde gaten verdwijnen direct in de p^+-laag.

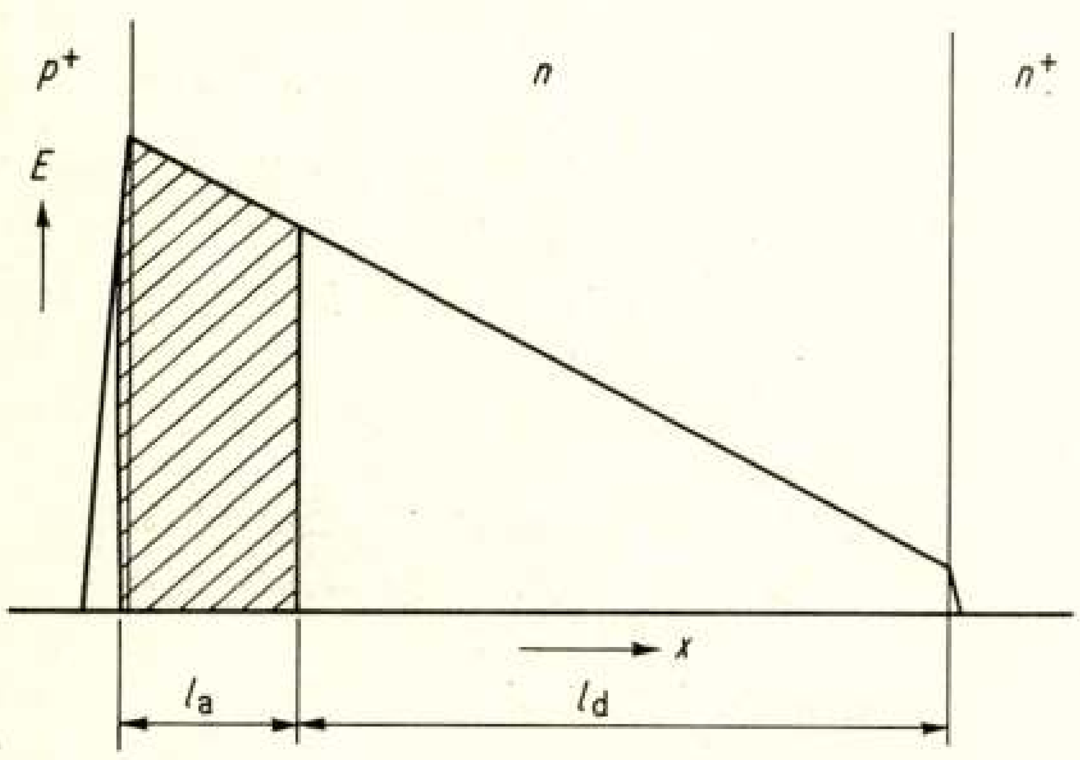


Fig. 1. Schematische voorstelling van het profiel van het elektrische veld in een lawine-looptijd-diode met asymmetrische abrupte p-n-overgang. In het gearceerd aangegeven doorslaggebied ter lengte l_a is de veldsterkte van de orde van 200 ... 500 kV/cm (silicium). Tussen opeenvolgende botsingen met het kristalrooster kunnen vrije ladingdragers in dit gebied voldoende energie uit het veld opnemen om stootionisatie te bewerkstelligen. In het driftgebied ter lengte l_d blijft de veldsterkte groter dan 10 ... 20 kV/cm.

Een boven de gelijkspanning aanwezige wisselspanning moduleert de in de doorslag gegenereerde hoeveelheid elektronen. In het als driftgebied fungerende gedeelte van de n-laag ondergaat ten gevolge van looptijd de gemoduleerde elektronenstroom een continue fasedraaiing ten opzichte van het wisselveld. Voor voldoend grote looptijd resulteert daaruit een negatieve weerstand. Heeft de modulatie van de elektronenstroom een tijdvertraging gelijk aan een kwart periode ten opzichte van het wisselveld, dan treedt negatieve weerstand op voor elke waarde van de looptijd. Dit geval van inductieve stroomgeneratie is karakteristiek voor de lawine-looptijd-diode.

Is $V_a(t)$ de boven de gelijkspanning aanwezige wisselspanning over het doorslaggebied, en is daar $\alpha(E)$ de van het elektrische veld afhankelijke ionisatiesnelheid, d.i. het aantal gat-elektronparen dat één ladingdrager per eenheid van lengte doet ontstaan, dan geldt voor de in het driftgebied geïnjecteerde convectiestroom $J_{ca}(t)$ bij benadering:

$$J_{ca}(t) = J_0 \exp\left\{\frac{2\alpha'_c}{\tau_a}\int_0^t V_a(\tau)\,d\tau\right\} \tag{1}$$

Hier is α'_c (>0) de bij het aangelegde dc-veld geëvalueerde afgeleide van α naar het elektrische veld en τ_a is de tijd benodigd voor het doorlopen van het doorslaggebied. De dc-instelstroom $J_{c0} = \langle J_{ca}(t)\rangle$ wijkt voor willekeurig signaalniveau $V_a(t)$ af van de stroom J_0; wordt de diode gevoed uit een constante-stroombron, dan is J_{c0} een onafhankelijk variabele. Uit vgl. (1) blijkt het inductieve karakter van de stroomgeneratie; voor kleine waarden van $V_a(t)$ kan dit karakter worden voorgesteld door een zelfinductie $L_a = \tau_a/(2\alpha'_c J_{c0})$.

De niet-lineariteit van de stroomgeneratie volgens vgl. (1) compliceert het gedrag van de lawine-looptijd-oscillator in belangrijke mate. Ten behoeve van het inzicht in de werking en het verkrijgen van ontwerpcriteria wordt in het volgende een analyse beschouwd van het niet-lineaire gedrag van de diode plus aangekoppeld microgolfcircuit. De vanwege het aangelegde dc-veld in het driftgebied geïnjecteerde convectiestroom $J_{ca}(t)$ gedraagt zich daar als een niet-dispersieve golf:

$$J_{cd}(x, t) = J_{cd}(0, t-x/v_d) = J_{ca}(t-x/v_d)$$

Hier is $J_{cd}(x, t)$ de convectiestroom in het driftgebied en v_d de veldverzadigde driftsnelheid van de ladingdragers. De wisselspanning over het driftgebied zij $V_d(t)$; de totale wisselspanning over de diode is dan $V_a(t) + V_d(t)$. Behalve de convectiestroom is ook de verschuivingsstroom continu op de overgang van doorslaggebied en driftgebied; in het eendimensionaal onderstelde model is de som van convectiestroom en verschuivingsstroom in het driftgebied divergentievrij.

Voor cyclische werking van de oscillator kunnen alle tijdafhankelijke grootheden worden voorgesteld door een fourierreeks. Zo schrijven we voor $J_{ca}(t)$:

$$J_{ca}(t) = J_{c0}\left\{1 + \sum_{m=1}^{\infty} (\beta_m e^{im\omega t} + \text{c.c.})\right\} \tag{2}$$

en voor $V_a(t)$:

$$V_a(t) = \sum_{m=1}^{\infty} (V_{am} e^{im\omega t} + \text{c.c.}) \tag{3}$$

Hier is i de imaginaire eenheid en de notatie c.c. duidt op de complex geconjugeerde term. De stroomamplitudes β_m worden bepaald door substitutie van vgl. (3) in vgl. (1); ieder van de stroomamplitudes β_m hangt dan (niet-lineair) af van alle spanningen V_{am}. Van belang is dat de stroom $J_{ca}(t)$ positief definiet is; daaruit volgt dat $|\beta_m| \leqq 1$. Is nu C_a de capaciteit van het doorslaggebied, dan wordt de $m$de harmonische component van de totale stroom op de overgang van doorslaggebied en driftgebied gegeven door

$$J_{tot,\,m} = \beta_m J_{c0} + im\omega C_a V_{am} \tag{4}$$

De impedantie van de diode voor de frequentie $m\omega$, $Z_{diode,\,m} = (V_{am} + V_{dm})/J_{tot,\,m}$, kan dan op grond van het aangenomen model worden berekend als

$$Z_{diode,\,m} = \frac{1}{im\omega C}\,\frac{1}{1+\beta_m J_{c0}/im\omega C_a V_{am}} + \frac{1}{im\omega C_d}\,\frac{1-(1-e^{-im\theta})/im\theta}{1+im\omega C_a V_{am}/\beta_m J_{c0}} \tag{5}$$

Hier is C_d de capaciteit van het driftgebied, C is de totale capaciteit van de diode volgens $C = C_a C_d/(C_a + C_d)$, en θ is de loophoek voor de frequentie ω volgens $\theta = \omega l_d/v_d$. Het looptijdeffect wordt uitgedrukt door de tweede term in vgl. (5). Behalve door de geometrische parameters C, C_d en θ, wordt het impedantiegedrag van de diode geheel bepaald door de verhoudingen $im\omega C_a V_{am}/\beta_m J_{c0}$, dus door de verhouding van verschuivingsstroom en convectiestroom in het generatiegebied. Is de generatie op de frequentie ω inductief, dan geldt $\beta_1 = -iV_{a1}|\beta_1/V_{a1}|$ en vertoont $Z_{diode,1}$ een parallelresonantie voor de frequentie waar $|\omega C_a V_{a1}| = |\beta_1 J_{c0}|$. De diode-impedantie $Z_{diode,1}$ is dan in principe beneden de resonantiefrequentie inductief en passief, boven de resonantiefrequentie capacitief en actief.

De werking van de oscillator wordt bepaald door de voorwaarde $Z_{diode,\,m} + Z_{em} = 0$ voor alle m, waarin Z_{em} de impedantie van het lineair en passief onderstelde externe circuit voor de frequentie $m\omega$ voorstelt. De diode-impedanties $Z_{diode,\,m}$ hangen af van de instelstroom en het signaalniveau, terwijl de impedanties Z_{em} niet afhangen van deze grootheden. Een zelfstartende oscillator voor bedrijf op een vooraf gespecificeerde frequentie en instelstroom vereist daarom in het algemeen een complex, zo niet onpraktisch, microgolfcircuit. Een bijzondere mogelijkheid bestaat nu indien voor waarden van J_{c0} boven de start-oscillatiestroom de spanningen V_{am} continu aanliggende waarden kunnen aannemen zodanig dat de verhoudingen $im\omega C_a V_{am}/\beta_m J_{c0}$ constant blijven. In dat geval blijven de diode-impedanties $Z_{diode,\,m}$ en dus de oscillatiefrequentie constant bij variërende J_{c0}, en wordt een welgedefinieerd ontwerp van het aan de diode te koppelen microgolfcircuit mogelijk. Van bijkomend belang is dat dan tevens parasitaire frequentiemodulatie ten gevolge van fluctuaties in J_{c0} vermeden wordt.

Om de genoemde mogelijkheid te onderzoeken, schrijven we de vergelijking $Z_{diode,\,m} + Z_{em} = 0$ met voordeel als een vergelijking voor de verhouding $im\omega C_a V_{am}/\beta_m J_{c0}$:

$$im\omega C_a V_{am}/\beta_m J_{c0} + \rho_m = 0 \tag{6}$$

Hier is:

$$\rho_m = (Z_{em} + Z_{dm})/(Z_{em} + 1/im\omega C)$$

en

$$Z_{dm} = (1/im\omega C_d)\cdot\{1-(1-e^{-im\theta})/im\theta\}$$

Zoals reeds opgemerkt is de stroomgeneratie voor voldoend lage waarden van $V_a(t)$ inductief, dus $\beta_1 = -iV_{a1}|\beta_1/V_{a1}|$. Oscillatie is voor lage signaalniveaus derhalve slechts mogelijk indien ρ_1 reëel is, hetgeen de afstemvoorwaarde

$$\text{Im}\, Z_{e1} = \frac{1}{\omega C} + (R_{s1} + \text{Re}\, Z_{e1})\left\{\cot(\theta/2) + \frac{C_d}{C_a}\,\frac{\theta/2}{\sin^2(\theta/2)}\right\} \tag{7}$$

oplevert. Hierin is Im Z_{e1} het reactieve deel van Z_{e1}, en R_{s1} + Re Z_{e1} het resistieve deel van Z_{e1} opgesplitst in een parasitair deel R_{s1} en een nuttig deel Re Z_{e1}. Voor reële waarden van ρ_1 geldt

$$\rho_1 = 1 + Q_1 \sin^2(\theta/2)/(\theta/2), \tag{8}$$

waarin $Q_1 = \{\omega C_d(R_{s1} + \text{Re}\, Z_{e1})\}^{-1}$ een maat is voor de kwaliteit van het oscillatorcircuit op de frequentie ω. Daar voor een passief extern circuit $\rho_1 > 1$, zien we dat aan vgl. (6) slechts voldaan kan worden voor frequenties boven de resonantiefrequentie, terwijl voor gegeven Q_1 de instelstroom een voldoend hoge waarde moet hebben om de oscillatie inderdaad te doen starten.

Een oplossing als hierboven voorgesteld blijkt nu voor een willekeurig signaalniveau niet mogelijk te zijn indien de stroomamplitudes β_m in essentie slechts bepaald worden door de grondgolf-spanning V_{a1}. Naar orde van analytische en ook in de praktijk realiseerbare gecompliceerdheid, onderzoeken we derhalve de mogelijkheid waarbij de stroomamplitudes β_m in essentie bepaald worden door spanningen V_{ak} op twee frequenties, ω en $n\omega$. De spanningen V_{ak} voor $k \neq 1,n$ zullen daartoe klein blijven ten opzichte van V_{a1} en V_{an}; volgens vgl. (6) vereist dit dat de coëfficiënten ρ_k voor $k \neq 1,n$ klein

gemaakt worden ten opzichte van ρ_1 en ρ_n. Met de spanningsnormalisatie $\bar{v}_m \equiv 2i\alpha_c' V_{am}/m\omega\tau_a$ worden dan de relevante oscillatorvergelijkingen:

$$-\beta_1/\bar{v}_1 = J_{c0,\,\text{start}\,\omega}/J_{c0}$$
$$-\beta_n/\bar{v}_n = (n^2\rho_1/\rho_n)\,(J_{c0,\,\text{start}\,\omega}/J_{c0}) \tag{9}$$

waarin $J_{c0,\,\text{start}\,\omega} \equiv (\varepsilon A/2\alpha_c' v_d)(\omega^2/\rho_1)$ die waarde van J_{c0} voorstelt waarvoor de klein-signaal-impedantie van het oscillatorcircuit voor de frequentie ω verdwijnt. De grootheid A is de oppervlakte van de dwarsdoorsnede van de diode.

De stroomamplitudes β_m worden dan voorts bepaald door de fourierontwikkeling van

$$\exp\{-(\bar{v}_1^{\,i\omega t} + \bar{v}_1{}^*e^{-i\omega t})\} \times \exp\{-(\bar{v}_n e^{in\omega t} + \bar{v}_n{}^* e^{-in\omega t})\}.$$

Daar slechts relatieve fasen van belang zijn stellen we $\bar{v}_1$ positief reëel; de complexe spanning $\bar{v}_n$ schrijven we als produkt van grootte en fasefactor: $\bar{v}_n \rightarrow v_n e^{i\psi_n}$. De gevraagde fourierontwikkeling vinden we als een som van mengprodukten van de door de eerste exponentiële functie gegenereerde harmonischen van de frequentie ω en de door de tweede exponentiële functie gegenereerde harmonischen van de frequentie $n\omega$. We vinden dan:

$$\beta_m = \frac{(-1)^m \sum\limits_{p=-\infty}^{+\infty} (-1)^{p(n+1)} I_p(2v_n)\, I_{pn-m}(2v_1)\, e^{ip\psi_n}}{\sum\limits_{p=-\infty}^{+\infty} (-1)^{p(n+1)} I_p(2v_n)\, I_{pn}(2v_1)\, e^{ip\psi_n}} \tag{10}$$

De functies $I_\nu(x)$ zijn de gemodificeerde Besselfuncties van de eerste soort, de orde ν en het argument x. Onmiddellijk volgt uit vgl. (10) dat β_1 voor willekeurig signaalniveau slechts een constante fase ten opzichte van v_1 kan behouden indien òf $\psi_n = 0$, òf $\psi_n = \pi$.

In dat geval wordt tevens β_n reëel en vereist oplossing van vgl. (9) een reële waarde van ρ_n. Deze eis levert voor de frequentie $n\omega$ een afstemvoorwaarde overeenkomstig die voor de frequentie ω, dus vgl. (7) voor de frequentie $n\omega$. De bijbehorende reële waarde van ρ_n wordt gevonden als vgl. (8) voor de frequentie $n\omega$.

Een belangrijke te bepalen grootheid is het door de oscillator op de frequentie ω geleverde nuttige vermogen. Door met vgl. (6) de totale stroom volgens vgl. (4) uit te drukken in $\beta_m J_{c0}$, vinden we dat voor reële ρ_1:

$$P_1 = \frac{1}{2}\left[2\,J_{c0}\,Q_1\,\frac{\sin^2(\theta/2)}{\theta/2}\right]^2 \beta_1\beta_1^*\ \text{Re}\ Z_{el} \tag{11}$$

Slechts de stroomamplitude β_1 is hierin de nog nader te bepalen grootheid. Daar $|\beta_1| \leqq 1$, levert vgl. (11) direct het maximaal haalbare uitgangsvermogen. De optimale loophoek wordt gegeven door $\theta = 0{,}74\pi$. Opgemerkt zij dat het maximaal haalbare uitgangsvermogen niet slechts gelimiteerd wordt door de eindige haalbare waarde van het produkt $Q_1{}^2$ Re Z_{e1}, maar tevens en vooral ook door de eindige toelaatbare waarde van de hoogfrequent spanningszwaai over de diode. In fig. 2 is uitgezet het door numerieke oplossing van vgln. (9) en (10) verkregen relatieve uitgangsvermogen $p_{1\,\text{rel}} \equiv (\beta_1 J_{c0}/J_{c0,\,\text{start}\,\omega})^2 = v_1{}^2$ als functie van $J_{c0}/J_{c0,\,\text{start}\,\omega}$ met ρ_n/ρ_1 als parameter voor een tweefrequentiecombinatie met $n = 2$. Als opvallende eigenschap zien we in fig. 2 voor iedere waarde van ρ_2/ρ_1 het bestaan van twee modi van oscillatie: een modus met hoog vermogen en een modus met laag vermogen. Deze twee modi van oscillatie komen respectievelijk overeen met $\psi_2 = \pi$ en $\psi_2 = 0$, hetgeen

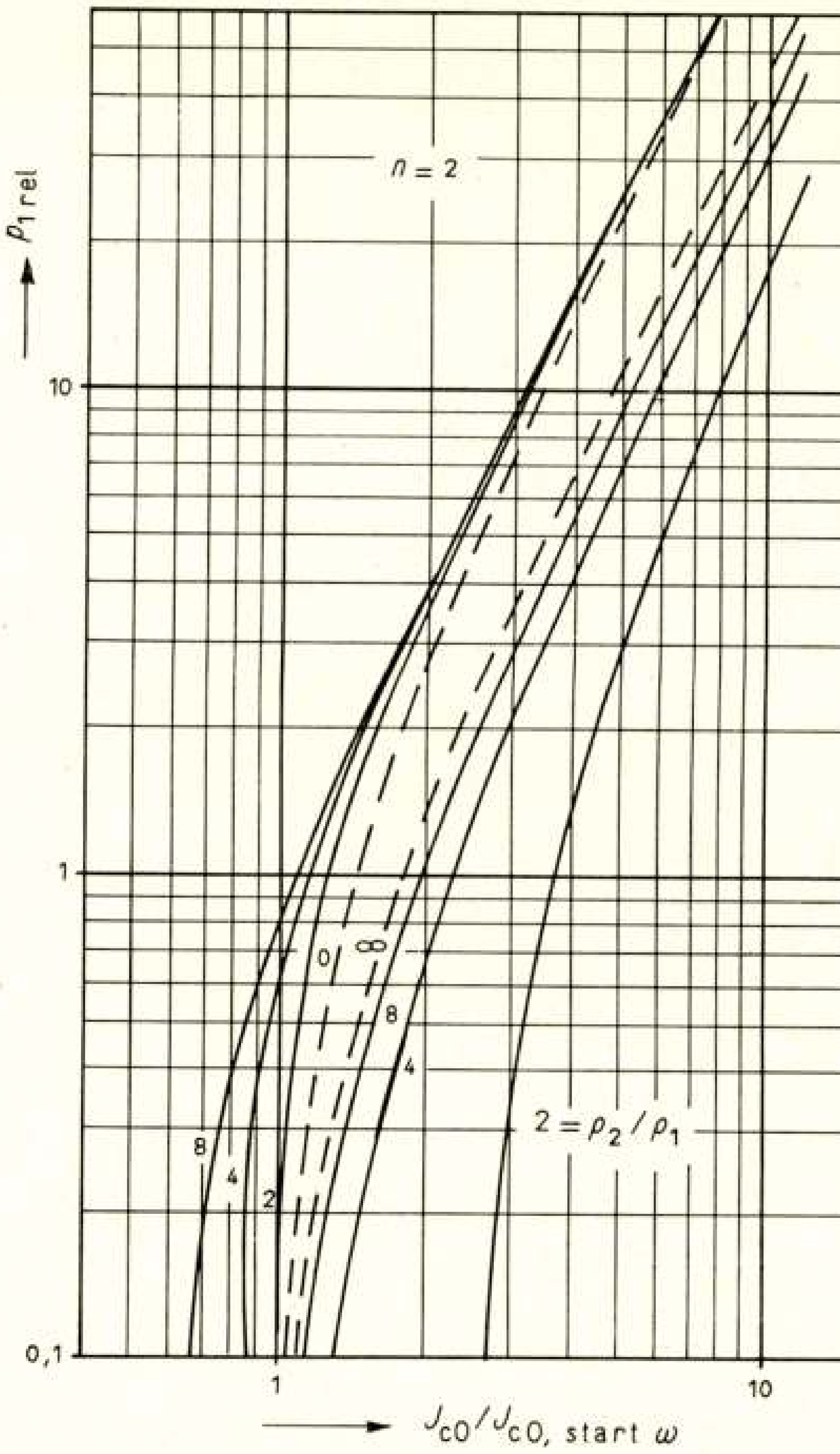


Fig. 2. Relatieve waarden van het op de frequentie ω geleverde nuttige vermogen als functie van de relatieve instelstroom bij gelijktijdig bedrijf op de frequenties ω en 2ω. De parameter ρ_2/ρ_1 is een maat voor de relatieve kwaliteit van het oscillatorcircuit op de frequenties ω en 2ω.

zich laat verklaren door het feit dat de oscillaties op de frequenties ω en 2ω, hoewel gekoppeld, ten opzichte van elkaar nog twee verschillende fasen kunnen hebben. De start-oscillatie-instelstroom kan verschillen van $J_{c0,\,\text{start}\,\omega}$ en verschilt voor de twee modi van oscillatie, behalve indien $\rho_2/\rho_1 = 4$. Ingeval $\rho_2/\rho_1 = 4$ komen de twee vermogenscurven voor $\rho_{1\,\text{rel}} \rightarrow 0$ samen bij $J_{c0}/J_{c0,\,\text{start}\,\omega} = 1$; de modus met hoog vermogen kan dan slechts bereikt worden via een sprong in oscillatie-amplitude. De modus met hoog vermogen heeft de laagste waarde van start-oscillatie-instelstroom, hetgeen praktische realisatie van deze modus vereenvoudigt. Het hoogste vermogen wordt voor beide modi verkregen voor de grootste waarden van ρ_2/ρ_1; daartoe dient de kwaliteitsfactor Q_2 van het oscillatorcircuit voor de frequentie 2ω hoog gekozen te worden. Opgemerkt zij dat tweefrequentiebedrijf voor waarden van $n > 2$ minder efficiënt blijkt dan bedrijf met $n = 2$. Voorts vormt een uiteindelijk oscillator-ontwerp nog een compromis voor wat betreft toelaatbare hoogfrequent spanningszwaai en voor wat betreft thermisch toelaatbare waarde van de dc instelstroom.

Voor een meer volledige beschrijving van de hier gegeven theorie en haar consequenties zij verwezen naar [2] en [3], waar tevens verdere verwijzingen naar de theorie van lawine-looptijd-oscillatoren gevonden kunnen worden. Opgemerkt zij nog dat met in het Natuurkundig Laboratorium in de groep van dr. D. de Nobel gemaakte dioden en constructie van de oscillator volgens de hier gegeven theorie, een nuttig vermogen van 1,75 W

in continu bedrijf op 5 GHz met een conversie-rendement van 9% kon worden bereikt.

Referenties

[1] W. T. READ, JR.: A proposed high-frequency negative resistance diode, Bell Sys. Techn. J., Vol. 37, pp. 401 ... 446, March 1958.

[2] K. MOUTHAAN: Nonlinear analysis of the avalanche-transit-time oscillator. I.E.E.E. Trans. Electron. Devices, Vol. ED-16, pp. 935 ... 945, November 1969. Nonlinear characteristics and two-frequency operation of the avalanche transit-time oscillator. Philips Res. Repts., Vol. 25, pp. 33 ... 68, February 1970.

[3] ———— and H. P. M. RIJPERT: Second-harmonic tuning of the avalanche transit-time oscillator. Proc. I.E.E.E. (Correspondence), Vol. 57, pp. 1449-1450, August 1969.

621.372.8

De toepassing van drie gekoppelde transmissielijnen

door ir. R. Blommendaal, Christiaan Huygenslaboratorium N.V., Noordwijk

Dedicated to Professor J. P. Schouten on the occasion of his retirement.



Summary: *The application of three coupled transmission lines.*
The theory of three coupled waveguides is outlined and applied to three waveguide rings bent in the *E*-plane. The input and output of the central ring are connected. This configuration can be used as a rotary joint if a gap is cut at the centre line of the middle ring. Expressions are given for the transmission as a function of rotation and frequency.

1. Inleiding

De theorie van de richtkoppeling tussen twee transmissielijnen is zeer goed bekend. In de literatuur is een overvloed aan stof te vinden over koppeling van golfgeleiders, coaxiale lijnen en 'stripline', waarbij zowel de continue als de discrete koppeling wordt behandeld. De golven in de secundaire lijn hebben steeds een faseverschil van 90° met die in de hoofdlijn en de amplitudes kunnen gevonden worden als functie van het koppelmechanisme en de lengte van de gekoppelde lijnen.

Het algemene geval van gelijktijdige en continue koppeling tussen een willekeurig aantal transmissielijnen is ook behandeld, maar slechts weinig bekend. Enerzijds komt dit omdat de oplossingen snel ingewikkeld worden en niet kunnen worden afgeleid uit de theorie van twee transmissielijnen, anderzijds omdat er weinig aantrekkelijke toepassingen bekend zijn die werk op dit gebied zouden kunnen stimuleren.

De theorie van drie gekoppelde golfgeleiders zal beknopt worden beschreven en worden toegepast op een geval waarbij de golfgeleiders in het *E*-vlak tot ringen gebogen zijn. Ingang en uitgang van de middelste golfpijp worden daarbij met elkaar verbonden. Deze configuratie is met een kleine wijziging bruikbaar als draaikoppeling indien de middelste ring op de neutrale lijn gedeeld wordt. Uitdrukkingen zullen worden gegeven voor de transmissie als functie van de frequentie en de stand van in- en uitgang.

2. Drie gekoppelde lijnen

In fig. 1 zijn drie gekoppelde lijnen geschetst. Gedacht wordt aan rechthoekige golfgeleiders met golfvoortplantingen in de TE_{01}-mode, die via de smalle zijde gekoppeld zijn. Voorts wordt aangenomen dat de koppeling continu verdeeld is, de golven zich alleen in de positieve *x*-richting voortplanten en dat de tijdafhankelijkheid $\exp(j\omega t)$ is. De volgende vergelijkingen kunnen dan worden opgesteld:

$$\begin{aligned}
\frac{dE_1}{dx} &= -j\beta E_1 - jc\,(E_1 + E_2)\\
\frac{dE_2}{dx} &= -j\beta E_2 - jc\,(E_1 + 2E_2 + E_3)\\
\frac{dE_3}{dx} &= -j\beta E_3 - jc\,(E_2 + E_3)
\end{aligned} \tag{1}$$

waarin: β = voortplantingsconstante
c = koppelfactor

Oplossingen van de vorm $E_n = \exp(-jrx)$ kunnen worden gezocht en blijken te zijn: $r_1 = \beta$, $r_2 = \beta + c$ en $r_3 = \beta + 3c$.

Met behulp van de randvoorwaarden: $E_1(0) = 1$, $E_2(0) = 0$, $E_3(0) = 0$ en $E_1(0) = 0$, $E_2(0) = 1$ en $E_3(0) = 0$ kunnen de integratieconstanten worden bepaald en daarmede de coëfficiënten van de strooimatrix:

$$\begin{bmatrix} E_1(x) \\ E_2(x) \\ E_3(x) \end{bmatrix} = e^{-j\beta x} \begin{bmatrix} \dfrac{e^{-j3cx} + 3e^{-jcx} + 2}{6} & \dfrac{e^{-j3cx} - 1}{3} & \dfrac{e^{-j3cx} - 3e^{-jcx} + 2}{6} \\ \dfrac{e^{-j3cx} - 1}{3} & \dfrac{2e^{-j3cx} + 1}{3} & \dfrac{e^{-j3cx} - 1}{3} \\ \dfrac{e^{-j3cx} - 3e^{-jcx} + 2}{6} & \dfrac{e^{-j3cx} - 1}{3} & \dfrac{e^{-j3cx} + 3e^{-jcx} + 2}{6} \end{bmatrix} \begin{bmatrix} E_1(0) \\ E_2(0) \\ E_3(0) \end{bmatrix} \tag{2}$$

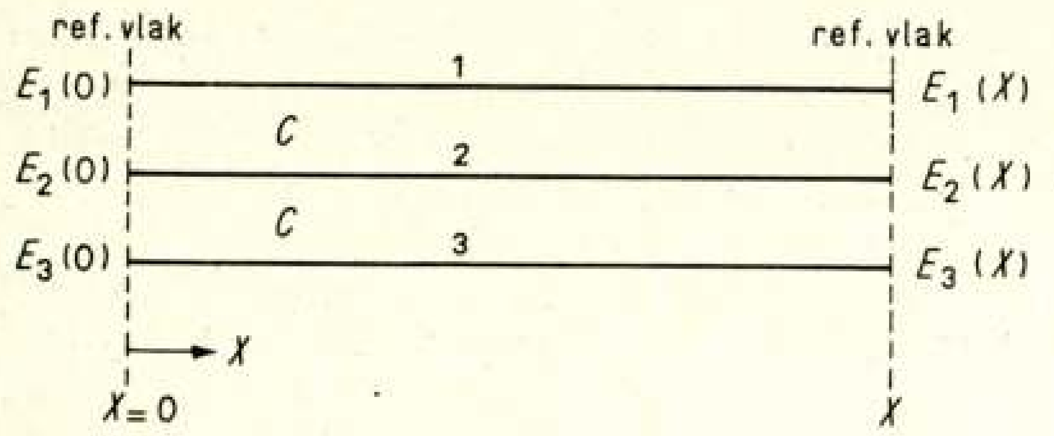


Fig. 1. Drie gekoppelde lijnen.

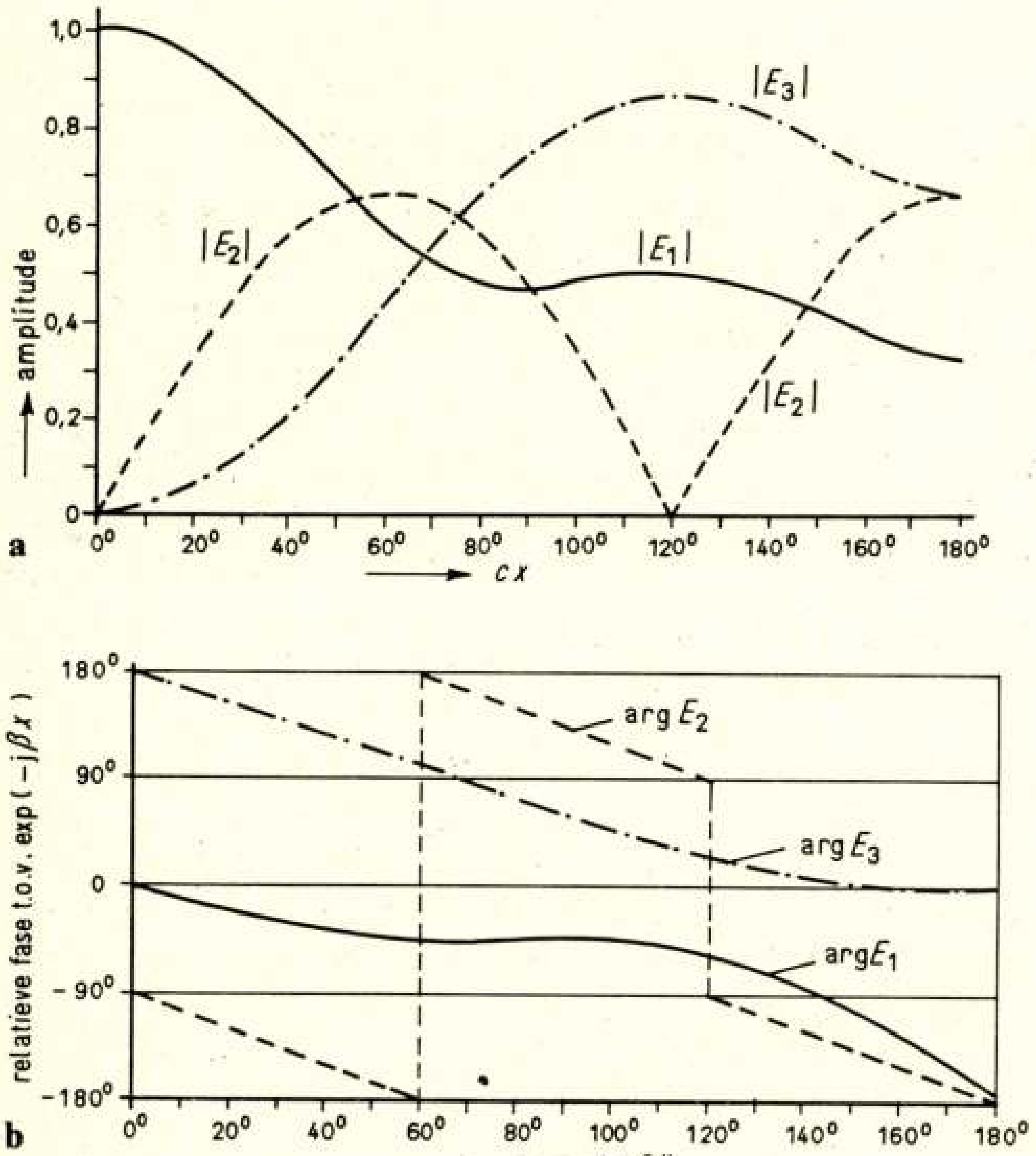


Fig. 2a, b. Amplitude en fase als functie van cx; $E_1(0) = 1$, $E_2(0) = E_3(0) = 0$.

Omdat gerichte koppeling is verondersteld, is hier volstaan met een 3 × 3-matrix. Het verloop van amplitude en fase als functie van cx is gegeven in fig. 2a resp. 2b voor het geval $E_1(0) = 1$ en $E_2(0) = E_3(0) = 0$. Voor geen enkele waarde van cx kan volledige vermogensoverdracht naar de lijnen 2 en 3 worden bereikt.

3. De ringvormige configuratie

De gevonden strooimatrix kan worden toegepast op de configuratie van fig. 3. Alle lijnen zijn gebogen met straal r, terwijl de middenring doorloopt. Het geheel kan worden opgevat als twee zespoorten ter lengte van $r\theta_1$ en $r\theta_2$. Indien alleen poort 1 met amplitude = 1 wordt gevoed en de uittredende golf in poort n wordt aangeduid met b_n, kunnen met behulp van (2) de volgende vergelijkingen worden opgesteld:

$$\begin{pmatrix} b_4 \\ b_5 \\ b_6 \end{pmatrix} = \begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \beta_1 & \delta_1 & \beta_1 \\ \gamma_1 & \beta_1 & \alpha_1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ b_{11} \\ b_{12} \end{pmatrix} \qquad (3)$$

$$\begin{pmatrix} b_{10} \\ b_{11} \\ b_{12} \end{pmatrix} = \begin{pmatrix} \alpha_2 & \beta_2 & \gamma_2 \\ \beta_2 & \delta_2 & \beta_2 \\ \gamma_2 & \beta_2 & \alpha_2 \end{pmatrix} \begin{pmatrix} b_4 \\ b_5 \\ 0 \end{pmatrix} \qquad (4)$$

met:

$$\left.\begin{array}{ll} \alpha_n = \dfrac{p_n^3 + 3p_n + 2}{6} R_n \; ; & \gamma_n = \dfrac{p_n^3 - 3p_n + 1}{6} R_n \\ \beta_n = \dfrac{p_n^3 - 1}{3} R_n & ; \; \delta_n = \dfrac{2p_n^3 + 1}{3} R_n \end{array}\right\} (n = 1 \text{ of } 2) \qquad (5)$$

$$\left.\begin{array}{ll} p_1 = e^{-jcr\theta_1} & ; \; p_2 = e^{-jcr(2\pi - \theta_1)} \\ R_1 = e^{-j\beta r\theta_1} & ; \; R_2 = e^{-j\beta r(2\pi - \theta_1)} \end{array}\right\} (0 < \theta_1 \leqq 2\pi)$$

Uit (3) en (4) volgt direct:

$$\left.\begin{array}{lll} (\beta_1\beta_2 + \gamma_1\gamma_2 - 1)\, b_4 + (\beta_1\delta_2 + \gamma_1\beta_2)\, b_5 & & = -\alpha_1 \\ (\delta_1\beta_2 + \beta_1\gamma_2)\, b_4 + (\delta_1\delta_2 + \beta_1\beta_2 - 1)\, b_5 & & = -\beta_1 \\ (\beta_1\beta_2 + \alpha_1\gamma_2)\, b_4 + (\beta_1\delta_2 + \alpha_1\beta_2)\, b_5 - b_6 & & = -\gamma_1 \end{array}\right\} \qquad (6)$$

Hieruit laat zich b_6 oplossen als functie van β, θ_1 en c. In dit korte bestek zou het te ver voeren dit algemene geval uit te werken. Beperken wij ons eerst tot het geval dat $\theta_1 = 2\pi$, dan geldt $\alpha_2 = \delta_2 = 1$ en $\beta_2 = \gamma_2 = 0$, zodat voor volledige vermogensoverdracht van 1 naar 6 moet gelden:

$$b_6 = \gamma_1 + \frac{\beta_1^{\,2}}{1 - \delta_1} = 1 \qquad (7)$$

Aan deze voorwaarde kan worden voldaan indien:

$$\left.\begin{array}{l} 2p_1^4 R_1 + 3p_1^3 R_1 - p_1^3 + p_1 R_1 - 3p - 2 = 0 \\ 2p_1^3 R_1 + R_1 - 3 \neq 0 \end{array}\right\} \qquad (8)$$

Mogelijke oplossingen:

$$\left.\begin{array}{l} p_1 = -1 \\ R_1 = 1 \end{array}\right\} , \quad \left.\begin{array}{l} p_1 = \mathrm{j} \\ R_1 = \mathrm{j} \end{array}\right\} \text{en} \quad \left.\begin{array}{l} p_1 = -\mathrm{j} \\ R_1 = -\mathrm{j} \end{array}\right\} \qquad (9)$$

De laatste oplossing kan worden gerealiseerd met de kortste lijnlengte en komt bovendien overeen met de voorwaarde voor volledige overkoppeling bij twee lijnen. Beschouwen we dit

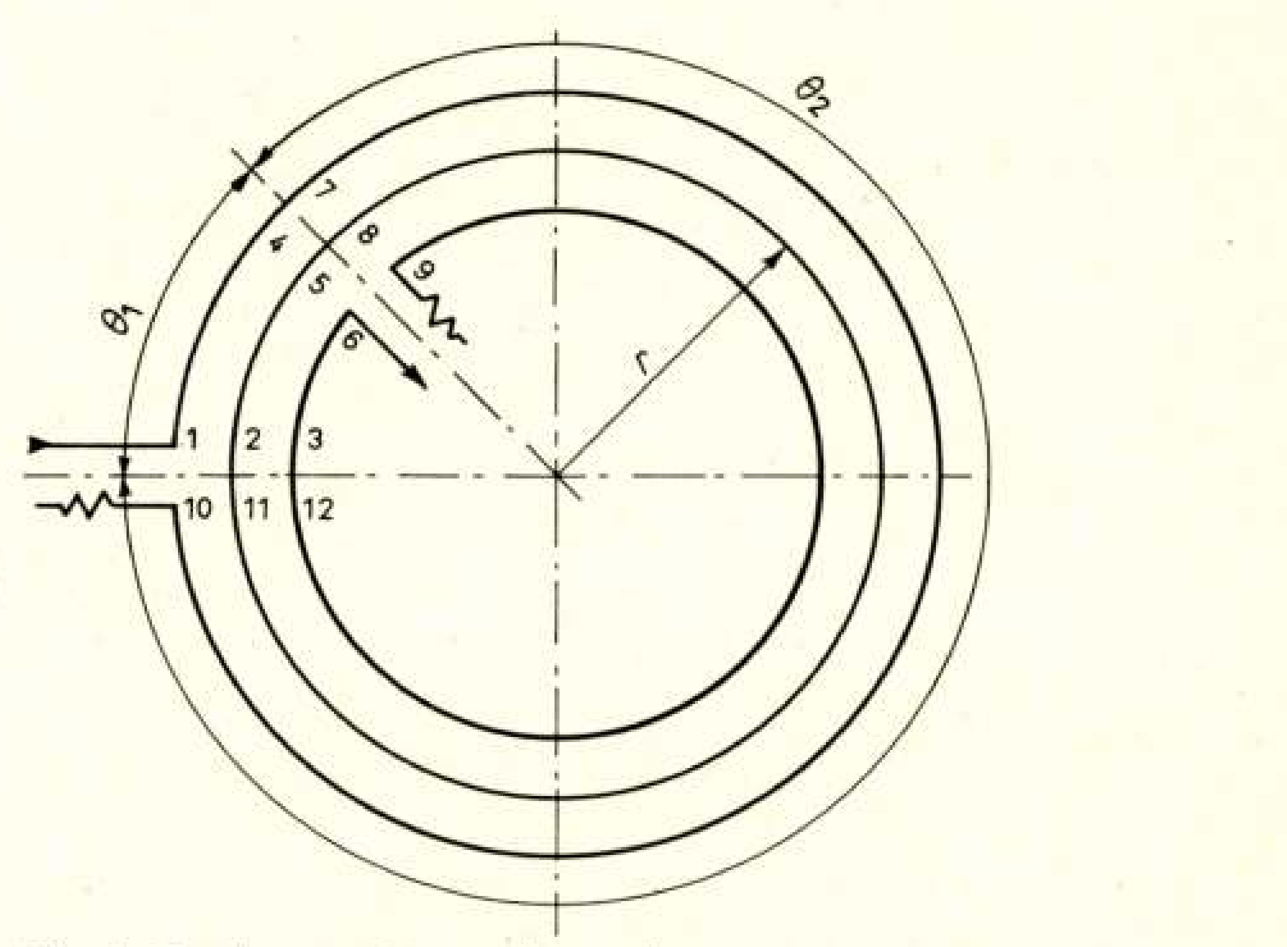


Fig. 3. De ringvormige configuratie.

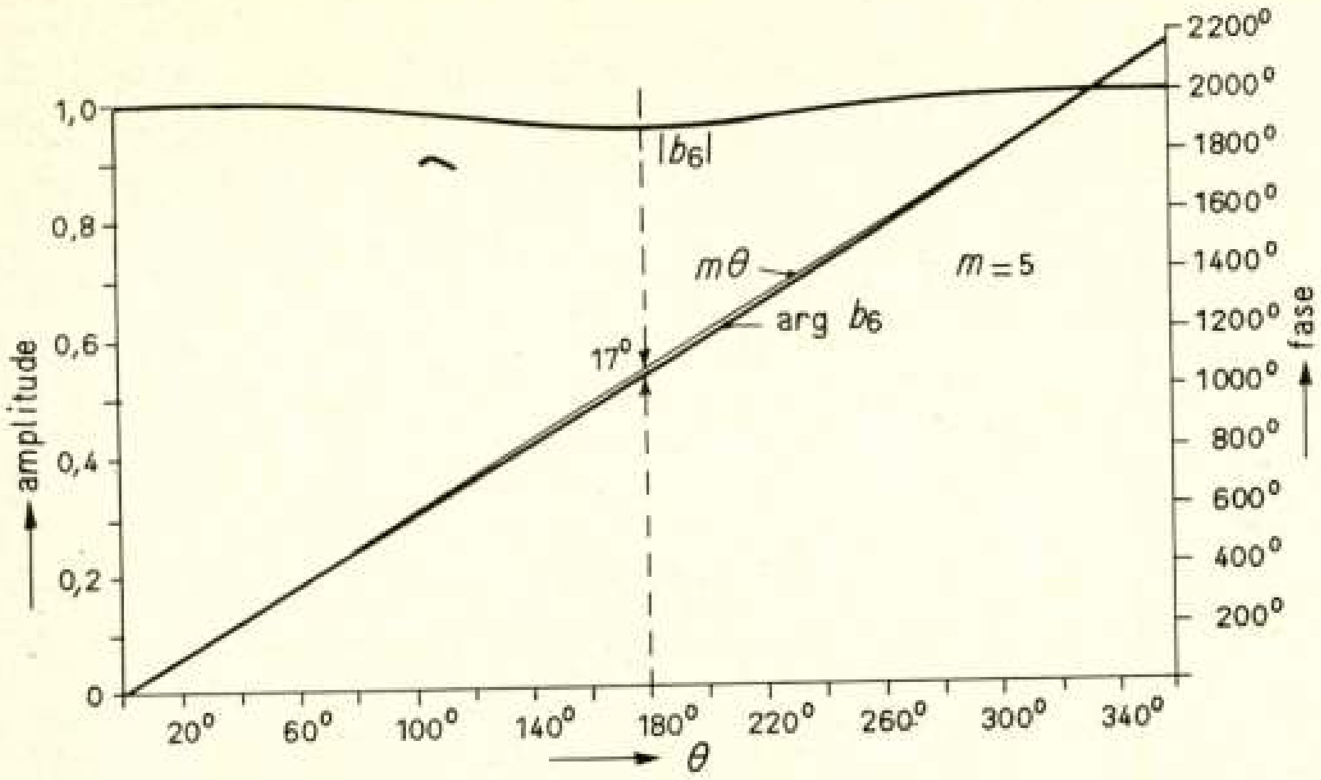


Fig. 4. De transmissie in amplitude en fase als functie van de stand θ.

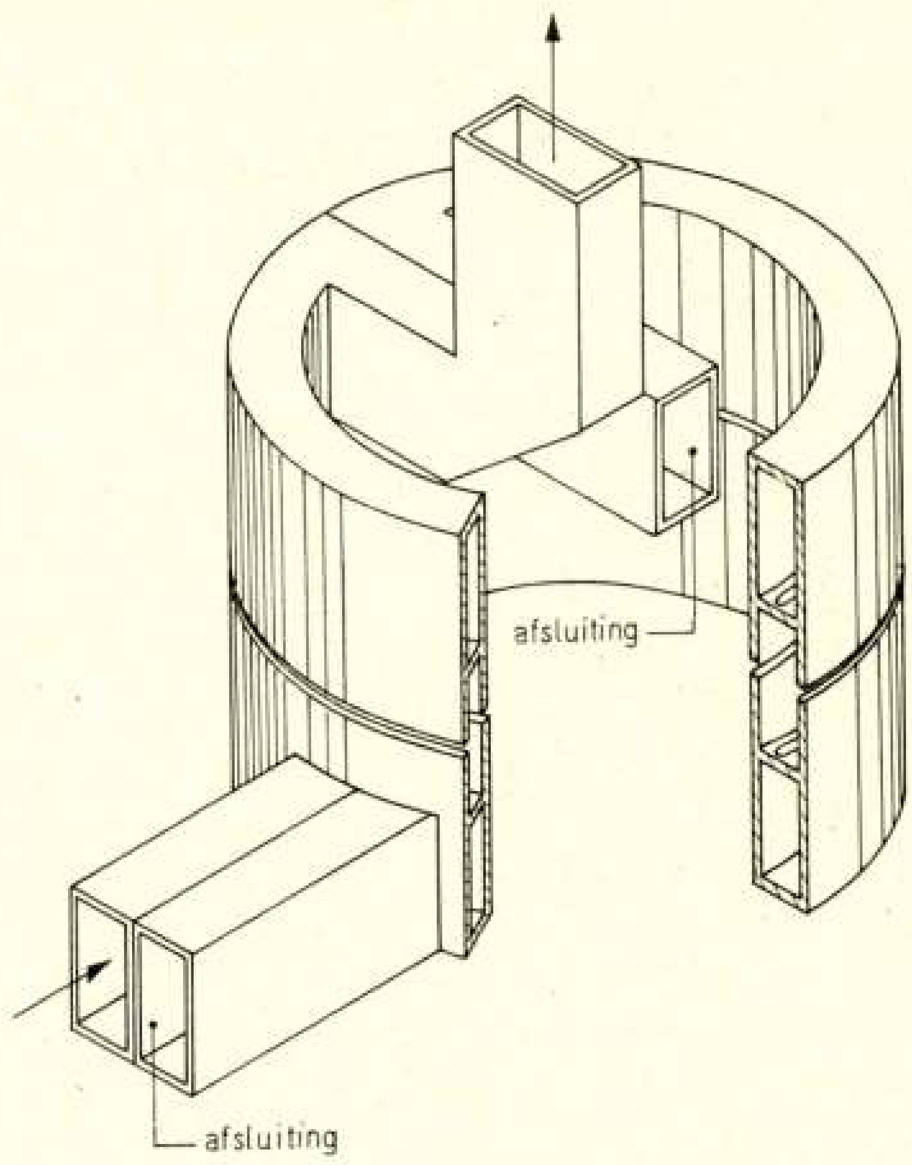


Fig. 5. Toepassing als draaikoppeling in rechthoekige golfpijp.

geval, waarbij $cr = 1/4$ en $\beta r = m + 1/4$ dan volgt uit (6) voor willekeurige $\theta_1 = \theta$:

$$b_6 = \frac{6e^{-j\frac{3\theta}{4}}\, e^{-j(\frac{1}{4}+m)\theta}}{2\,j\cos\frac{3\theta}{4} - 2\,j\sin\frac{3\theta}{4} + 3\,j\sin\frac{\theta}{2} + 6 - 2\,j} \quad (10)$$

waarin m geheel voor de hartfrequentie.

4. De toepassing

Het verloop van de transmissie van poort 1 naar poort 6 als functie van θ is gegeven in fig. 4 ($m = 5$). De amplitude is constant binnen $\pm$ 0,25 dB, terwijl de fase maximaal 17° van een lineaire toeneming afwijkt. Deze uitvoering is geschikt als draaikoppeling indien de middenring op de neutrale lijn gedeeld wordt zoals in fig. 5. Beide helften kunnen dan vrij gedraaid worden ten opzichte van elkaar. De variatie van de amplitude ten gevolge van het ronddraaien kan nog verminderd worden indien men de voortplantingsconstante van de middenring wijzigt. Dit kan gedaan worden door de breedte van de golfpijp te veranderen, wat door de aanwezigheid van deling eenvoudig is. De theorie van dit geval is ook te vinden op de bovenbeschreven wijze, maar is zeer bewerkelijk.

De ringvormige draaikoppeling heeft het voordeel dat zij op eenvoudige wijze uitgebreid kan worden tot meerdere kanalen, door identieke exemplaren te 'stapelen'. Bij de gebruikelijke coaxiale draaikoppeling is dit niet mogelijk, al kan men een systeem met twee kanalen realiseren door de binnengeleider ook als coaxiale lijn uit te voeren. Het ruime gat in de ringvormige draaikoppeling maakt het bijv. mogelijk alle naar boven lopende lijnen door de gaten te voeren, en de naar beneden lopende aan de buitenzijde te houden. Naast het nadeel van verandering van de fase met de stand heeft dit soort draaikoppelingen een beperkte bandbreedte: 5 à 10%. Het gedrag als functie van de frequentie kan eveneens uit de boven gegeven uitdrukkingen worden berekend. Er treden in de 0°-stand stopbanden op nabij frequenties waar de weglengte door de middenring een oneven aantal halve golflengten is. Een praktische uitvoering voor frequenties van 8800 ... 9200 MHz heeft de bruikbaarheid van de rekenwijze aangetoond.

Literatuur

S. Boronski: A Multichannel Waveguide Rotating Joint. The Microwave Journal, June 1965, pp. 102 ... 105.

J. P. Shelton Jr.: Multiple-Line Directional Couplers. IRE – Nat. Conv. Rec. Pt. I, 1957, pp. 254 ... 262.

621.396.67

Modern and classical antenna design techniques

by dr. ir. M. S. Afifi, Delft University of Technology; now with Directorate for Research and Engineering of Page Communications Engineers, Inc., Washington D.C.



Summary: A short review is given of the history of antenna design techniques. Some viewpoints are explained for modern techniques in comparison with the classical ones. More attention is given to frequency-independent antennas and the necessity of finding simple mathematical concepts in order to comply with the requirements of evolving physical pictures.

1. Introduction

Electromagnetic waves are controlled by physical rules which are formulated in a compacted form by Maxwell's equations. The antenna is a device for generating these waves with given specifications which are determined by propagation require-

ments. This antenna must also comply with other requirements which insure complete transmission of electromagnetic energy from the generator to free space, when transmitting, or from free space to the receiver, when receiving. The problem is to find an adequate configuration and material for this device in order to verify the previously mentioned requirements. In general there is no routine method for solving this problem. Since the early days of electromagnetic waves attention has been directed to applying metallic conductors. Recently, lossless dielectric materials were introduced [1] and lately, semiconductors [2, 3] and superconductors [4] are being successfully applied in design of active and passive antennas.

Disregarding the problem of material choice, it is very difficult to find by analysis a metallic antenna configuration which complies with given requirements of impedance and directivity. Development in this area is a matter of experience and convenience. It is striking to mention that when the computer was invented – with its great storage capacity – some workers in this field thought that it could replace the requirement of experimental work as it is able to give the needed information, just by solving Maxwell's equations with boundary conditions on the antenna surface [5]. In this statement though, no attention was given to how to choose this surface. The antenna problem however, is such that neither the current nor the surface configuration are known. Classical antenna design techniques, or antenna theory, is based on choice of the configuration and then finding the current distribution, the impedance, the radiation pattern and the frequency characteristics. The difficulty of choice of the configuration to suit required radiation patterns has stimulated modern developments of antenna arrays [6], where a lot of flexibility exists in arranging and feeding different elements.

The object of this paper is to throw some light on usual antenna design techniques and a trial to find an adequate way which can be followed in the design considerations. As the subject is too comprehensive to be adequately condensed into a small article, few points are carefully selected to give clear representation, without loss of generality, applicable to most antenna types.

2. General considerations

Developments of antenna theory are always mentioned in connection with solutions of Maxwell's equations for different known antenna configurations, and a classical discussion of antennas is usually based on some simplified assumptions which do not fit very exactly into a rigorous theory. This began in the early days of electromagnetic wave theory with the discussion of Hertz's experiments. Hertz is the founder of antenna theory; after few years from the publication of Maxwell's treatise on electromagnetic theory he has solved the first, the simplest and the most basic antenna problem of a small capacitor antenna (oscillating doublet or Hertzian dipole). After the year 1900 [7, 8] a number of very important theoretical papers were published on spherical, ellipsoidal, cylindrical and biconical antennas. These gave better understanding of the phenomenon. For time dependence of $e^{j\omega t}$ the solution of Maxwell's equations (with the boundary conditions on the surface of the antenna) is found possible only for certain values of ω (the so-called *proper values*) and these are derived as complex quantities. From these complex quantities the Q-factor of the antenna can be defined by analogy to normal resonance circuits. If $\omega = \omega_r + j\omega_i$, from the free vibration of the antenna Q_f will be $\omega_r/2\omega_i$. It has been learned also that the Q-factor of spherical antennas for different modes are very small, some of which are even zero, which corresponds to damping without oscillations. This certainly showed that spherical antennas are typical examples of wide-band antennas.

3. Linear antennas

The current along a thin ellipsoidal antenna and long prolate spheroids shows practically sinusoidal distribution when the length of the antenna is half a wavelength, which corresponds to resonance. The distribution deviates from this sinusoidal form away from resonance. The sinusoidal form is the base for what is called the *conventional theory of antennas*. Linear antennas of half a wavelength are assumed, with a good degree of approximation, to carry currents having this form.

More accurate determination of the form of this distribution is required for longer antennas. Trials to achieve this requirement began by Hallén for getting an approximate solution for the z-component of the vector potential, A_z, due to currents flowing in the longitudinal direction of a linear antenna parallel to the z-axis. The differential equation for this vector potential is given bij

$$\left[\frac{\partial^2}{\partial z^2} + k^2\right] A_z(z) = -j\frac{k}{\omega} V_0\,\delta(z)$$

where $k = 2\pi/\lambda$, λ being the wavelength of electromagnetic radiation; $\delta(z)$ is the Dirac delta-function and V_0 is the applied voltage at the central gap of the linear antenna, whose length to diameter ratio is assumed to be large. The solution of this differential equation yields the following integral equation for the current distribution [9]

$$\int_{-h}^{h} I_z(z')\frac{e^{-jkr}}{r}\,dz' = \left\{ j4\pi/(\mu_0/\varepsilon_0)^{\frac{1}{2}} \right\} \left\{ C\cos kz + \frac{1}{2}V_0 \sin k|z| \right\}$$

where h is half the antenna height, $r = \{(z-z')^2 + a^2\}^{\frac{1}{2}}$, a is the radius of the conductor, $ka \ll 1$, $a \ll h$ and C is a constant which can be determined from the boundary conditions. King [9] and his colleagues proposed different methods for solving this integral equation with some approximation. They could develop lately various practical solutions for this integral equation. Their three-term method gave good approximation to practical measurements. They could also extend their five-term method, which is concerned with two parallel antennas to include any number of parallel antennas. A nice survey of these efforts is given by King, Wu and Chen in a recent publication [10].

The efforts directed to the solution of the previously mentioned integral equation in nearly 80 years did not contribute as much to the design techniques of linear antennas, as the contribution of the conventional theory of antennas which is based on the concept of the sinusoidal form of the current distribution. This simple concept and others which are introduced into the antenna theory early after 1900 could lead to a lot of developments in antenna design techniques. One of these developments, and the most important one, is the Yagi-Uda antenna, which has been in use since 1926. Some practical and theoretical design techniques of this antenna [11] appeared in 1954. More accurate calculations for the current distribution on this antenna have not been available till the last few years when many workers began to formulate the problem systematically

by the availability of the previously mentioned expressions for multielement systems [12].

The bandwidth of a linear antenna is determined by its thickness. This phenomenon has been learnt from the characteristics of the prolate spheroid. It has been proved that the thinner the prolate spheroid in question, the higher the equivalent Q and the smaller the frequency bandwidth. Very thick antennas approach, in their configuration, the previously mentioned sphere, which has been proved to be a typical example of a broadband antenna.

4. Frequency-independent antennas

These are known since 1957 when Rumsey broke the antenna bandwidth barrier with his angle concept. This concept was already known to Rumsey earlier than 1957, but he delayed its publication until Dyson demonstrated in September 1957, in the *antenna laboratory* of Illinois University, that the log-spiral is indeed frequency-independent. The representation of the field generated by a frequency-independent antenna is easy if Maxwell's equations are written in a form for which the differentiations are performed with respect to the normalized dimensions to the wavelength. When these differentiations are represented by rot′, Maxwell's equations reduce to [13]:

$\mathrm{rot}' \underline{E} \propto \underline{H}$ and $\mathrm{rot}' \underline{H} \propto \underline{E}$,

where the constant of proportionality is independent of the frequency. This means that if the tangential field on the surface of the antenna is fixed, and the dimensions of the antenna scale with the wavelength, the antenna will be frequency-independent. The next problem now is how to find the antenna which satisfies this specification and which suits the required application. At this point the ability of the human brain to understand the physical phenomenon and find out the adequate configuration of the antenna comes into action. In 1957 DuHamel and Isbell invented the log-periodic antenna which has found more wide applications than the original antenna described by Rumsey. It is interesting to mention also that some designs of the log-periodic antenna which were first proposed by Turner did not work (these designs were not published but remain in the history of development of these antennas). In 1959 Isbell's first log-periodic antenna could work. Lo [14] grasped the concept of an *active region* which moves with frequency and saw the necessity of the twisted feeder or reversal of alternate dipoles. This concept and the universal relation which relates the antenna impedance to the impedance of its complementary structure ($4Z_1Z_2 = Z_0^2$) are the most powerful tools in this field. The analysis of these antennas (some workers in this field call it 'theoretical work'; to my idea the theoretical work includes also the stage of finding out the shape of the antenna structure as well as its material) lagged in time, until experimental results were available for interpretation, because a rigorous solution of Maxwell's equations is prohibitively difficult for their complicated structure. Mittra and his colleagues published some papers in the last few years on these analyses [15, 16].

5. Microwave and superdirective antennas

In the field of microwave antennas the developing procedure did not differ much from that of the previously mentioned types of antennas except that the idea of using metallic reflectors and lenses as antennas is taken from the known optical devices. The most useful concept which is of wide application in this field is the fourier transform relation between the radiation pattern of an antenna and its aperture field distribution [17]. The author succeeded to find mathematical concepts which were helpful in simplifying the design techniques of parabolic and other modified types of reflector antennas. The most simple concept is what is called *the edge relation* for the illumination function at the reflector edge [18, p. 46].

In the field of superdirective antennas many efforts are in process since the publication of Bouwkamp and de Bruijn [19] in which they have shown that there is no limit for increasing the directivity of the antenna by choice of an idealized current distribution. There are practical limitations for this phenomenon because of the large amplitudes and high rates of phase variation of these currents on the antenna structure [20]. This unfavourable condition arises not only when it is attempted to increase the directive gain excessively, but also when it is required to closely approach a desired nonrealizable directive pattern or to realize unjustified decrease in the sidelobe level [21, p. 1788]. An example of the phenomenon of superdirectivity which is associated with excessive decrease in the side lobe level is given in [18, p. 92].

6. Conclusions

From the previous discussion it is easy to see that the antenna design techniques present the following three stages:

1. The stage of finding an adequate configuration of the antenna to satisfy the radiation requirements and to suit the field of application. In this stage idealized situations can be studied, from which simplified assumptions and conclusions may be obtained. This stage is characterized by fascinating simplification of the electromagnetic theory in combination with observation which give physical insight into the problem. Successful results in this stage depend on our stored knowledge as to the behaviour of fields and currents. This can help in correlating the simple facts of electromagnetic phenomena and constructing a pattern for the current distribution and its supporting structure in order to arrive at the adequate configuration. It is interesting to mention here the words of Rumsey in the introduction of his book on frequency-independent antennas [13]: 'Research in this area has uncovered many striking phenomena which remain unexplained, and has generated solutions of Maxwell's equations which have yet to be observed'.

2. In most of the cases the proposed construction in the previous stage may be complicated or it may not work. Therefore it is preferred preliminarily to test the possible practical applicability of these antennas before proceeding to the stage of analysis. This helps, firstly to gain more experience for interpretation and guidance of the assumptions for mathematical solutions of the complicated configurations, and secondly to avoid the risk of time waste if, for some reason or another which could not be seen in the first stage, the originally proposed constructions are not practically realizable. This stage, and the previous stage, are especially essential for antennas in which unconventional materials, such as semiconductors, may be used.

In addition to this conclusion it is interesting to mention these words of Levi [22]: 'Measurement is fundamental to all science, and the formulation of concepts used in such measurement is itself an important part of the science'.

3. Experience from the previous two stages is essential for performing useful mathematical analysis. The object of such analysis is to help the designer in evolving his original physical picture and bringing order to the multiplicity of facts obtained in the first two stages. When, sometimes, a quantitative solution of the problem is hard to find, those physical pictures often give qualitative answers. Maxwell's equations themselves express a physical point of view. At this point it is interesting to mention the definition of 'practical solution' as given by King [9, Chapt. 9]: 'Solutions are called practical if they combine reasonable quantitative accuracy with formal simplicity in a physically meaningful representation. Such solutions can be used, on the one hand, to solve engineering problems and, on the other hand, to illuminate the underlying physical principles in their relation to the actual situation'.

As the simplicity of the mathematical formulation of antenna problems contrasts sharply with the complexity of its solution, the stage of developing a useful practical solution is the longest and difficult stage, especially when the designer is guided by others' physical concepts. Some designers, however, use concepts which are inconsistent with Maxwell's equations and yet with this aid some approximate results have been obtained. Such a concept is that the power losses due to radiation produce the same effect on transmission of waves along conducting wires as losses due to dissipation. We hope that the use of modern computers with their modern useful applications will give a hand to help shortening this stage of antenna design techniques.

Finally, it is clear to see that the theory of antennas with its three closely related stages is a typical example of the close relation between science and technology, the relation described by Bondi [23] as that of 'the chicken and the egg': you cannot have the one without the other. This is actually the physical nature of science.

References

[1] Kiely, D. G.: Dielectric Aerials. Methuen, London, 1953.

[2] Meinke, H. H.: Aktive Antennen. N.T.Z., **19**, 12, pp. 697 ... 704, 1966.

[3] Meinke, H. H.: Tunnel diode integrated with microwave antenna systems. The Radio and Electronic Engineer, pp. 76 ... 80, February 1966.

[4] Walker, G. B. and C. R. Haden: Superconducting antennas. Journ. of Appl. Physics, **40**, 5, p. 2035, 1969.

[5] Tanner, R. L. and M. G. Andreasen: Numerical solution of electromagnetic problems. IEEE Spectrum, pp. 53 ... 61, September 1967.

[6] Radio Science: Special number on antenna arrays. **3** (new series), 5, May 1968.

[7] Brillouin, L.: Antennae for ultra-high frequencies, wideband antennae. Electrical Communication, **21**, pp. 257 ... 281, 1942-1944 and **22**, p. 11-39, 1944-1945.

[8] Schelkunoff, S. A.: Advanced antenna theory. New York, John Wiley and Sons, Inc. 1952.

[9] King, R. W. P.: The linear antenna – eighty years of progress. Proc. IEEE, **55**, 1, pp. 2 ... 16, January 1967.

[10] Collin, R. E. and F. J. Zucker: Antenna theory, part I. McGraw-Hill Book Company, 1969.

[11] Uda, C. and Y. Mushiake: Yagi-Uda antenna. Sendai, Japan, 1954.

[12] Thiele, G. A.: Analysis of Yagi-Uda antennas. IEEE Trans. AP-17, 1, January 1969.

[13] Rumsey, V. H.: Frequency independent antennas. Academic Press, 1966.

[14] Lo, Y. T.: TV receiving antennas. Antenna Engineering Handbook, H. Yasik, Editor, Chapt. 24, McGraw Hill, New York, 1961.

[15] Mittra, R. and K. E. Jones: Nonuniform transmission lines with application to log-periodic antennas. Proceedings of the National Electronics Conference, **20**, pp. 23 ... 28, 1964

[16] Jones, K. E. and P. E. Mayes: Continuously scaled transmission lines with application to log-periodic antennas. IEEE Trans. AP-17, **1**, pp. 2 ... 9, January 1969.

[17] Silver, S.: Microwave antenna theory and design. Radiation Laboratory Series No. 12, McGraw-Hill Co. Inc., 1949.

[18] Afifi, M. S.: Scattered radiation from microwave antennas and the design of paraboloid-plane reflector antenna, Thesis, T.H. Delft, July 1967.

[19] Bouwkamp, C. J. and N. G. de Bruijn: The problem of optimum antenna current distribution. Philips Research Reports, **1**, pp. 135 ... 158, 1945-1946.

[20] Hansen, R. C.: Microwave scanning antennas – Vol. I, (Chapt. 1, Sect. 3). Academic Press, 1964.

[21] Bakhrakh, L. D., A. A. Pistol 'kors and Ya. N. Fel'd: Developments in antenna technology. Radio Eng. and Electronic Phys., **12**, 11, pp. 1787 ... 1813, 1968.

[22] Levi, L.: Applied Optics, Chapt. I, p. 1. John Wiley & Sons, Inc., New York/London/Sydney, 1968.

[23] Bondi, H.: Assumption and myth in physical theory. Cambridge University Press, 1967.

778.3:535.3



Asymptotic theory of holographic image formation

by dr. ir. H. J. Frankena, Delft University of Technology, Department of Electrical Engineering

Summary: Using the exact formulas of scalar diffraction theory, the field behind an illuminated hologram is investigated. The results are expanded asymptotically as $k \to \infty$, leading to expressions which describe well-known phenomena that usually are deduced from approximate expressions valid only for restricted classes of configurations.

1. Introduction

In considerations regarding the formation of images through the illumination of a hologram, authors always use, from the beginning of their calculations, approximations based on the smallness of the optical wavelength as compared to the spatial distances in the configuration and from the smallness of the object, the hologram and the images as compared to the distances between them. The present account intends to deliver a more exact theory on the same subject, thereby only em-

phasizing the smallness of the wavelength with respect to the various distances in the configuration. Our aim is to show the application of certain methods rather than to give the most general proofs. Therefore, we often shall restrict the account to simple cases.

The electromagnetic field will be represented by the complex scalar function $U(x,y,z;\omega)$ (see, e.g., [1]), where we assume the field to vary harmonically in time with the factor $\exp(-i\omega t)$; this factor is suppressed in all subsequent formulas. The function U is a solution of Helmholtz's equation

$$\nabla^2 U + k^2 U = 0, \tag{1.1}$$

where $k = \omega/c$ is the wave number and c = velocity of propagation of the optical waves.

We follow the well-known procedure in which we firstly calculate the intensity $|U|^2$ at the place of the hologram during its formation and then, assuming the transparency of the hologram after development to be proportional to this intensity, calculate the field created by a wave after its transmission through the hologram during the reconstruction phase.

The field U during the formation phase consists of the superposition of the field U_{ob}, coming from the illuminated object and of the field U_{ref}, usually called the reference wave. From a known object, illuminated by a field U of angular frequency ω_1 and placed in a homogeneous and isotropic medium we have the contribution [2]:

$$U_{ob}(x,y,z;\omega_1) = \iint_{object} \Big\{ \partial U(x_{ob}, y_{ob}, z_{ob}; \omega_1)/\partial n + \\ - U(x_{ob}, y_{ob}, z_{ob}; \omega_1)(ik_1 - R_{ob}^{-1})\partial R_{ob}/\partial n \Big\} \frac{\exp(ik_1 R_{ob})}{4\pi R_{ob}} \, d\,S_{ob} \tag{1.2}$$

where $\partial/\partial n$ denotes the derivative at the surface of the object in the direction of the inward normal and k_1 is the wave number in this situation. Here, R_{ob} stands for

$$R_{ob} = \{(x-x_{ob})^2 + (y-y_{ob})^2 + (z-z_{ob})^2\}^{\frac{1}{2}} \geqq 0 \tag{1.3}$$

Choosing the medium to be lossless we know that k_1 is real. For the reference wave we insert a plane wave of the same time dependence as U_{ob}, travelling in a direction given by the phase vector $\vec{\beta}_1$. Its field U_{ref} is then

$$U_{ref}(x,y,z;\omega_1) = A_1 \exp(i\vec{\beta}_1 \cdot \vec{r}) \tag{1.4}$$

where

$$\vec{r} = x\vec{i}_x + y\vec{i}_y + z\vec{i}_z \tag{1.5}$$

and A_1 is a constant. We imagine the hologram to be situated in the plane $z = 0$ and the fields U_{ob} and U_{ref} are assumed to be incident from the half-space $z < 0$. The total intensity of light I incident upon the film is then

$$\begin{aligned} I(x,y) &= |U_{ob} + U_{ref}|^2_{z=0^-} = \\ &= |U_{ob}(x,y,0^-;\omega_1)|^2 + |U_{ref}(x,y,0^-;\omega_1)|^2 + \\ &\quad + U_{ob}(x,y,0^-;\omega_1)\, U^*_{ref}(x,y,0^-;\omega_1) + \\ &\quad + U^*_{ob}(x,y,0^-;\omega_1)\, U_{ref}(x,y,0^-;\omega_1) \end{aligned} \tag{1.6}$$

We suppose, in accordance with statements from specialists in this field, that the transparency t of the developed hologram is directly proportional to I. If, during reconstruction, a wave coming from $z < 0$ strikes the hologram, the field transmitted through the hologram will be

$$U(x,y,0^+;\omega_2) = t(x,y)\, U_{rec}(x,y,0;\omega_2) \\ (x, y \text{ on the hologram}), \tag{1.7}$$

where U_{rec} is the wave function describing that wave. We assume the angular frequency during reconstruction to be ω_2. Taking for U_{rec} again a plane wave, we have

$$U_{rec}(x,y,z;\omega_2) = A_2 \exp(i\vec{\beta}_2 \cdot \vec{r}) \tag{1.8}$$

where A_2 is a constant. With (1.7), the function $U(x,y,0^+;\omega_2)$ is, apart from a constant factor, equal to the product of the expressions at the right-hand sides of (1.6) and (1.8). At points in the half-space $z > 0$, the field U can be obtained by application of a representation theorem. We let the hologram occupy the rectangle $-a \leqq x \leqq a$, $-b \leqq y \leqq b$, $z=0$; outside this rectangle we take the factor t in (1.7) to be zero. Let the wave number in this situation be k_2; then, from the representation theorem for a half-space [3]

$$U(x_{im}, y_{im}, z_{im}; \omega_2) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} U(x,y,0;\omega_2) \times \\ \times \frac{\partial}{\partial z_{im}} \left\{ \frac{\exp(ik_2 R_{im})}{R_{im}} \right\} dx dy \tag{1.9}$$

we arrive with (1.5) ... (1.7) at

$$U(x_{im}, y_{im}, z_{im}; \omega_2) = c \int_{-a}^{a} \int_{-b}^{b} U_{rec}(x,y,0;\omega_2) \times \\ \times \frac{\partial}{\partial z_{im}} \left\{ \frac{\exp(ik_2 R_{im})}{R_{im}} \right\} [|U_{ob}|^2 + |U_{ref}|^2 + U_{ob} U^*_{ref} + \\ + U^*_{ob} U_{ref}] \, dx dy \qquad (z_{im} > 0) \tag{1.10}$$

In these formulas, R_{im} denotes

$$R_{im} = [(x_{im} - x)^2 + (y_{im} - y)^2 + z_{im}^2]^{\frac{1}{2}} \geqq 0 \tag{1.11}$$

c is a constant factor while the arguments of U_{ob} and U_{ref} in the expression between square brackets have been given in (1.6).

2. Discussion of the various contributions to the wave function during reconstruction

In order to discuss the result (1.10), we consider the contributions of each term between the square brackets separately.

a. Inserting (1.8), we obtain from the first term a contribution of the general form

$$U_1 = \frac{\partial}{\partial z_{im}} \int_{-a}^{a} \int_{-b}^{b} f_1(x,y) \exp(i\beta_{2,x}\, x + i\beta_{2,y}\, y + \\ + ik_2 R_{im})\, R_{im}^{-1} \, dx dy \tag{2.1}$$

where in this case $f_1(x,y)$ contains information concerning the amplitude of the field coming from the object but not concerning its phase. In many calculations, authors do neglect this term in comparison with the next term in (1.10). This procedure seems to be incorrect, as most experimentors try to obtain almost equal values of $|U_{ob}|^2$ and $|U_{ref}|^2$ at the hologram plane during the formation.

In view of the fact that k_2 has a large value for optical frequencies, we can expand the integral in (2.1) asymptotically

as $k_2 \to \infty$. Thereby, we assume $|U_{ob}|^2$ to vary slowly as compared with the exponential functions in (2.1) (both $\beta_{2,x}$ and $\beta_{2,y}$ contain a factor k_2). We apply the method of stationary phase [4]; for simplicity, we restrict the considerations to the first approximation in which only the contribution of the stationary point is taken into account. Introducing the direction cosines $\gamma_{2,x}$, $\gamma_{2,y}$ and $\gamma_{2,z}$ through

$$\beta_{2,x} = k_2 \gamma_{2,x}, \qquad \beta_{2,y} = k_2 \gamma_{2,y},$$
$$\gamma_{2,z} = \{ 1 - \gamma_{2,z}{}^2 - \gamma_{2,y}{}^2 \}^{\frac{1}{2}} > 0 \tag{2.2}$$

we obtain from (2.1) by subsequent application of the method of stationary phase on the integrations with respect to y and to x:

$$U_1 = F_1(x_{im}, y_{im}, z_{im}) \times$$
$$\times \exp\{ ik_2(\gamma_{2,x}\ x_{im} + \gamma_{2,y}\ y_{im} + \gamma_{2,z}\ z_{im})\} \quad (z_{im} > 0, k_2 \to \infty) \tag{2.3}$$

where F_1 is an algebraic function of the coordinates x_{im}, y_{im}, z_{im} in which a.o. the function f_1 from (2.1), calculated in the stationary point $x = x_{im} - \gamma_{2,x}|z_{im}/\gamma_{2,z}|$, $y = y_{im} - \gamma_{2,y}|z_{im}/\gamma_{2,z}|$, appears as a factor. We do not consider the special case $\gamma_{2,z} = 0$ (striking incidence on the hologram), which has no practical use. From this, we obtain a non-zero contribution for any x_{im}, y_{im}, z_{im} in the relevant half-space; it forms nowhere an image and acts as a spurious illumination, spoiling possible images created by the further contributions to U in (1.10). We remark that the asymptotic expansion can be extended to one giving contributions of higher-order terms too, such as the contributions from the end points of the intervals of integration. However, such an extension will not lead to a solution that yields a better understanding of the phenomenon described by U_1.

b. Considering the second term between square brackets in (1.10) we remark that, from (1.4), $|U_{ref}|^2$ is only a constant while $U_{rec}(x, y, z; \omega_2)$ is a solution of Helmholtz's equation with the same angular frequency ω_2 as occurs in the factor between braces in (1.10). Then, on account of the uniqueness of the contribution U_2 as given by

$$U_2 = c|A_r|^2 \int_{-a}^{a} \int_{-b}^{b} U_{rec}(x, y, 0; \omega_2) \frac{\partial}{\partial z_{im}} \left\{ \frac{\exp(ik_2 R_{im})}{R_{im}} \right\} dx dy \tag{2.4}$$

for a known distribution $U_{rec}(x, y, 0; \omega_2)$, the field U_2 is, apart from a constant factor, the plane wave U_{rec} from (1.8) which propagates through a rectangular aperture into the half-space $z_{im} > 0$.

Hence, this second term describes a contribution which consists of the reconstructing plane wave which travels through the hologram into the half-space $z_{im} > 0$, thereby undergoing a change in amplitude.

c. Inserting (1.2), (1.4) and (1.8) into the third term of (1.10), we arrive at the contribution U_3:

$$U_3 = \frac{\partial}{\partial z_{im}} \iint_{object} \left\{ \int_{-a}^{a} \int_{-b}^{b} f_3(x_{ob}, y_{ob}, z_{ob}; x, y, 0^-) R_{ob}^{-1} R_{im}^{-1} \times \right.$$
$$\times \exp(ik_1 R_{ob} + ik_2 \gamma_{2,x} x + ik_2 \gamma_{2,y} y + ik_2 R_{im} - ik_1 \gamma_{1,x} x -$$
$$\left. - ik_1 \gamma_{1,y} y) dx dy \right\} dS_{ob} \tag{2.5}$$

where $\gamma_{1,x}$ and $\gamma_{1,y}$ come from $\vec{\beta}_1$ in a way similar to (2.2).

Here, f_3 is the function between braces in (1.2) multiplied by a constant. For simplicity, we restrict the calculations to the case $k_1 = k_2 = k$ and, again, try to obtain from (2.5) an asymptotic expansion for U_3 as $k \to \infty$ by means of the method of stationary phase. Furthermore, we put $\gamma_{1,x} = \gamma_{2,x}$ and $\gamma_{1,y} = \gamma_{2,y}$ which means that U_{ref} and U_{rec} propagate in the same direction.

Writing the exponential function at the right-hand side of (2.5) in the form $\exp\{ikh_3(x, y)\}$, we know that in the stationary point $\partial h_3/\partial x = 0$ and $\partial h_3/\partial y = 0$ simultaneously, or

$$(x - x_{ob})/R_{ob} = (x_{im} - x)/R_{im}$$
$$(y - y_{ob})/R_{ob} = (y_{im} - y)/R_{im} \tag{2.6}$$

In order to visualize the first relation, we consider Fig. 1 and remark that in the stationary point we have $\sin(\alpha_{ob}) = \sin(\alpha_{im})$ or, in this case, $\alpha_{ob} = \alpha_{im}$ (in Fig. 1, the points P_{ob} and P_{im} have the co-ordinates x_{ob}, y_{ob}, z_{ob} and x_{im}, y_{im}, z_{im}, respectively). The set of points in the plane $z = 0$, satisfying this relation, is a curve of which in any case the point P (being the point at $z = 0$ of the straight line between P_{ob} and P_{im}), is a member. Analogously, the second relation (2.6) yields a set of points containing P. Hence, we may conclude that P is the stationary point; its co-ordinates are

$$(x_{ob}z_{im} - x_{im}z_{ob})/(z_{im} - z_{ob}), \quad (y_{ob}z_{im} - y_{im}z_{ob})/(z_{im} - z_{ob}), \quad 0$$

(recall that $z_{ob} < 0$). Further we remark that for $k_1 = k_2$ we have

$$h_3(x, y) = R_{ob} + R_{im} \tag{2.7}$$

which in the stationary point P is equal to the distance R_{obim} between the points P_{ob} and P_{im}.

After application of the method of stationary phase at the integrals with respect to x and y, we obtain from (2.5)

$$U_3 = \iint_{object} F_3(x_{ob}, y_{ob}, z_{ob}; x_{im}, y_{im}, z_{im}) \frac{\exp(ik R_{obim})}{R_{obim}} dS_{ob}$$
$$(k \to \infty) \tag{2.8}$$

in which F_3 is, apart from a constant factor, equal to the product of the expression between braces in (1.2) (which itself, provided

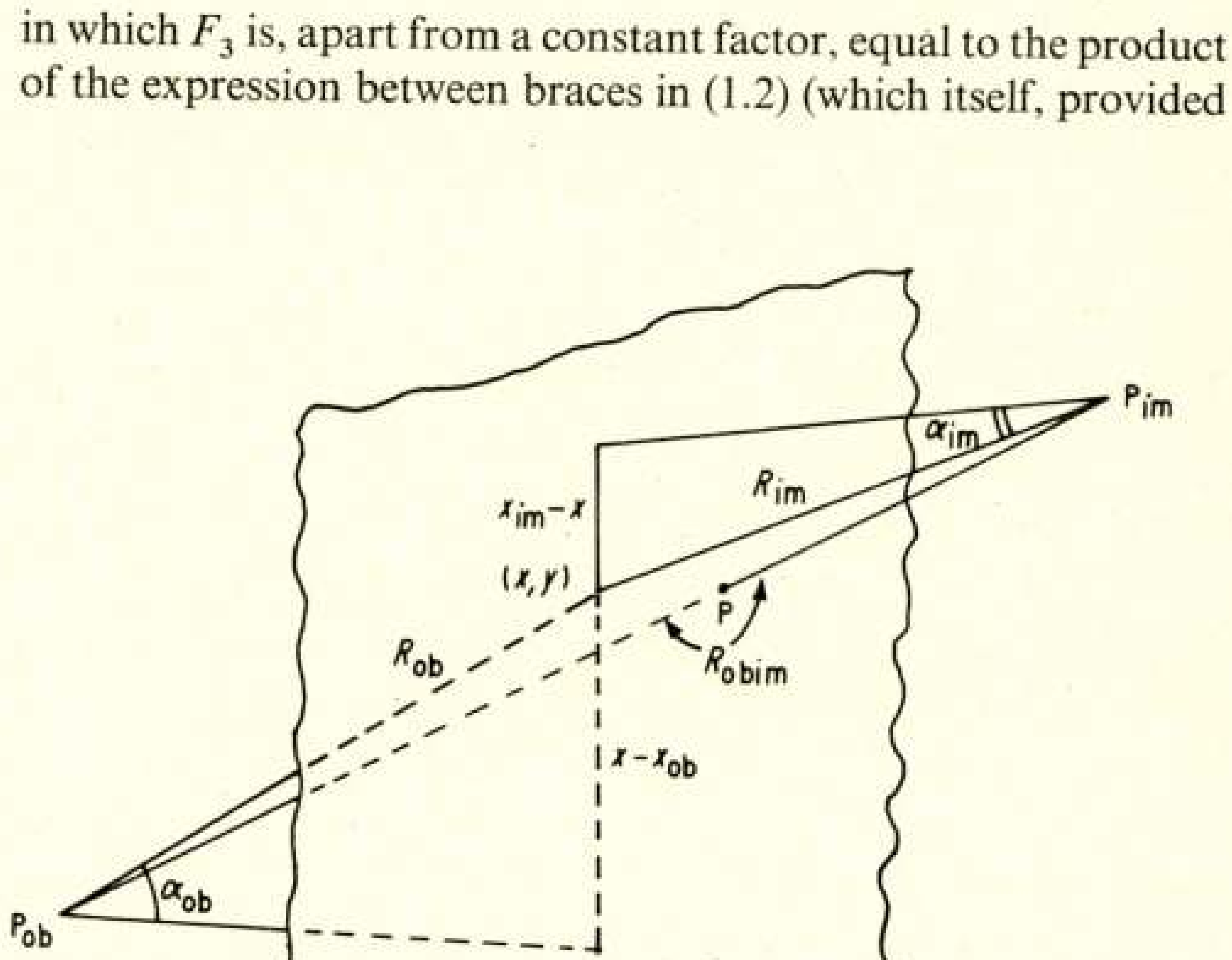


Fig. 1. The geometry from which the stationary point P is deduced.

that $R_{ob} \neq 0$ (object outside hologram plane), can be reduced for $k \to \infty$) and an algebraic function of the co-ordinates of both P_{ob} and P_{im}. This function, which is related to the second derivatives of h_3 in the stationary point, will be almost constant if z_{ob} varies only slightly over the object and if $|z_{ob}|$ is large as compared with $|x_{im}-x_{ob}|$ and $|y_{im}-y_{ob}|$. In that case, the right-hand side of (2.8) is proportional to the right-hand side of (1.2) with x, y, z replaced by x_{im}, y_{im}, z_{im}. This means that then the field U_3 for $z_{im} > 0$ is, apart from a constant factor, the analytic continuation in $z_{im} > 0$ of the field which was created by the object in its original position. It means the occurrence of a 'virtual image' at the same place as where the object was situated during the formation phase. If, however, $|z_{ob}|$ is not large enough, the algebraic factor in F_3 will cause deformation of this image.

The present considerations apply only to the case where the size and position of the hologram are such that the stationary point is inside the intervals $-a<x<a$, $-b<y<b$. In any other case, an asymptotic expansion will yield only the higher order effects due to the boundaries of the hologram and no virtual image can be seen. This is easily demonstrated in an experimental setup by choosing the position of an observer (x_{im}, y_{im}, z_{im} in the formulas) improperly.

d. The last term between square brackets in (1.10) yields the contribution

$$U_4 = \frac{\partial}{\partial z_{im}} \iint_{\text{object}} \left\{ \int_{-a}^{a}\int_{-b}^{b} f_4\,(x_{ob}, y_{ob}, z_{ob}; x, y, 0)\, R_{ob}^{-1}\, R_{im}^{-1}\, \times \right.$$

$$\left. \times \exp(-ik_1 R_{ob} + ik_2 \gamma_{2,x} x + ik_2 \gamma_{2,y} y + ik_2 R_{im} + ik_1 \gamma_{1,x} x + + ik_1 \gamma_{1,y} y)\, dx dy \right\} dS_{ob} \tag{2.9}$$

with the same notation as used in (2.5). The function f_4 differs, apart from constants, from f_3 in the complex conjugation of the expression between braces at the right-hand side of (1.2). Again, we put $k_1 = k_2 = k$ and we write the exponential factor in the integrand as $\exp\{ikh_4(x, y)\}$.

The considerations are now restricted to the simple case $\gamma_{1,x} = \gamma_{1,y} = \gamma_{2,x} = \gamma_{2,y} = 0$, i.e. the case that both U_{ref} and U_{rec} are travelling along the z-axis. Then we have

$$h_4\,(x, y) = -R_{ob} + R_{im} \tag{2.10}$$

The stationary point follows from

$$(x-x_{ob})/R_{ob} = -(x_{im}-x)/R_{im}$$
$$(y-y_{ob})/R_{ob} = -(y_{im}-y)/R_{im} \tag{2.11}$$

It is now simpler to proceed algebraically than to follow the geometrical observations of the preceding section.

From (2.11) we deduce (as $z_{ob} < 0$)

$$(x-x_{ob})/(x-x_{im}) = (y-y_{ob})/(y-y_{im}) = R_{ob}/R_{im} = -z_{ob}/z_{im} \tag{2.12}$$

From (2.10) we can write

$$h_4 = \pm\,\{R_{ob}^{\,2} + R_{im}^{\,2} - 2R_{ob}\, R_{im}\}^{\frac{1}{2}} \tag{2.13}$$

in which, in the stationary point, with (2.12)

$$R_{im}\, R_{ob} = R_{im}^{\,2}\,(x-x_{ob})/(x-x_{im}) =$$
$$= (x-x_{ob})\,(x-x_{im}) + (y-y_{ob})\,(y-y_{im}) - z_{ob}z_{im} \tag{2.14}$$

Insertion of (1.3) (with $z = 0$), (1.11) and (2.14) in (2.13) leads to

$$h_4 = \pm\{(x_{ob}-x_{im})^2 + (y_{ob}-y_{im})^2 + (z_{ob}+z_{im})^2\}^{\frac{1}{2}} = \pm R'_{obim} \tag{2.15}$$

in the stationary point, where R'_{obim} is the distance from the observation point x_{im}, y_{im}, z_{im} to the point x_{ob}, y_{ob}, $-z_{ob}$. The sign in (2.15) has to be chosen in accordance with (2.10), in which R_{ob} and R_{im} are both positive. Hence, in the domain $R_{im} \geqq R_{ob}$, we arrive at an asymptotic expansion for the integrations with respect to x and y in (2.9) of the form $\exp(ikR'_{obim})$ times an algebraic factor which is practically constant when $|z_{ob}|$ varies only slightly over the object, while $|z_{ob}| \gg |x_{im}-x_{ob}|$ and $|z_{ob}| \gg |y_{im}-y_{ob}|$. For $R_{im} < R_{ob}$, this expansion has $\exp(-ik\, R'_{obim})$ as the exponential factor. Now, we obtain for U_4.

$$U_4 = \iint_{\text{object}} F_4\,(x_{ob}, y_{ob}, z_{ob}; x_{im}, y_{im}, z_{im}) \times$$

$$\times \frac{\exp(\pm ik\, R'_{obim})}{R'_{obim}}\, dS_{ob} \qquad (k \to \infty) \tag{2.16}$$

For $R_{im} < R_{ob}$, this means waves converging to a domain found by mirroring the object with respect to the plane $z = 0$; for $R_{im} \geqq R_{ob}$ these waves are diverging from that domain. This means, physically, that a 'real image' is formed there. All remarks in section 2c. concerning the deformation of the image by the above-mentioned algebraic factor and concerning the non-existence of this image if the size and the place of the hologram are ill-chosen, apply here too.

References

[1] Born, M., and E. Wolf: Principles of Optics. Oxford, Pergamon Press, 1964, p. 387.

[2] Baker, B. B., and E. T. Copson: The mathematical theory of Huygens' Principle. Oxford, Clarendon Press, 1950, pp. 68-71.

[3] Ibid., p. 157.

[4] Erdelyi, A.: Asymptotic Expansions. Dover Publications Inc., 1956, p. 51.

717.56

Een gegeneraliseerde Schlömilch-reeks



door ir. H. W. F. van 't Groenewout, Rotterdam-Overschie

Dedicated to Professor J. P. Schouten on the occasion of his retirement

Summary: *A generalized Schlömilch series.*
An itemized survey is given of several infinite Bessel series which are closely related to the well-known Schlömilch series. The series studied were derived from an investigation of the quantization noise that goes with a pulse-code-modulation system.

1. Inleiding

Toen kort na de oorlog, mede onder invloed van het baanbrekende werk van Shannon op het gebied van de informatietheorie, belangstelling ontstond voor nieuwe modulatiesystemen, gaf Schouten een zeer oorspronkelijke methode aan om de vervorming te berekenen die bij pulscode-modulatie (PCM) in de vorm van quantiseringsruis optreedt. Daarbij werd gebruik gemaakt van de laplacetransformatie om de sprongen die ontstaan als het signaal achtereenvolgens de equidistante drempels overschrijdt, mathematisch weer te geven. De resultaten van deze studie werden gepubliceerd en naast een inzicht in het juiste gedrag van de modulatieprodukten werd een bevestiging gevonden van de grootte van de signaal-ruisverhouding van een dergelijk modulatiesysteem zoals die door anderen langs geheel verschillende weg was gevonden [1, 2].

Bij deze studie kwam het, kort samengevat, hier op neer dat de gezochte quantiseringsruis kon worden uitgedrukt als de fourierontwikkeling van de zaagtandfunctie waarbij als abscis niet de tijd zelve doch het ingangssignaal als functie van de tijd gekozen wordt. Dit leidt voor bijv. sinusvormige ingangssignalen tot uitdrukkingen van het type

$$\sum_{r=1}^{\infty} \frac{\sin\{2\pi r A \sin(\omega t)\}}{\pi r} \tag{1}$$

Een nadere uitwerking van deze reeksontwikkeling leert dat de amplitudes van de gezochte modulatiecomponenten gegeven worden door reeksen van het type

$$\sum_{r=1}^{\infty} a_r J_\nu(2\pi r z) \tag{2}$$

een oneindig voortlopende reeks dus van Besselfuncties van de eerste soort waarbij het argument van iedere term evenredig is met de rang van de term. Reeksen van dit type zijn eerder onderzocht door Schlömilch en anderen, doch voor het betrokken onderzoek bleek het noodzakelijk te zijn een verdergaande studie te maken van reeksen van een meer algemeen karakter en wel

$$\sum_{r=1}^{\infty} a_r J_\nu(2\pi r z)\, \phi(2\pi r y) \tag{3}$$

waarin $\phi(2\pi r y)$ een goniometrische functie voorstelt.

Gezien de bijzonder snelle en elegante wijze waarop men met de laplacetransformatie het doel had bereikt, lag het voor de hand om ook voor de sommering van deze bijzondere reeksen van dezelfde transformatiemethode gebruik te maken. De resultaten van de berekeningen langs deze weg en wel van reeksen van het type:

$$\sum_{r=1}^{\infty} \frac{J_{2n+\nu}(2\pi r z)}{(\pi r z)^\nu} \tag{4}$$

$$\sum_{r=1}^{\infty} (-)^r \frac{J_{2n+\nu}(2\pi r z)}{(\pi r z)^\nu} \tag{5}$$

$$\sum_{r=1}^{\infty} \frac{J_{2n+\nu}\{2\pi(2r-1)z\}}{\{\pi(2r-1)z\}^\nu} \tag{6}$$

werden in een intern rapport [3] vastgelegd en de toegepaste rekenmethode in een publikatie [1] in hoofdlijnen weergegeven.

In het onderstaande zullen wij ons bezighouden met de sommatie van de meer algemene reeks

$$\sum_{r=1}^{\infty} \frac{J_{2n+\nu}(2\pi r z)}{(\pi r z)^\nu} \cos(2\pi r y), \tag{7}$$

die behoort tot de categorie onder (3) genoemd en waarvan in wezen de reeksen (4), (5) en (6) als limietgevallen beschouwd kunnen worden voor respectievelijk $y = 0$ en $y = \frac{1}{2}$, terwijl de reeks (6) gevonden wordt door het verschil te nemen van (4) en (5). Bij deze berekening wordt om de mathematische aarzelingen te overwinnen, die nu eenmaal inherent zijn aan een rechtstreekse toepassing van de laplacetransformatie bij de termsgewijze integratie van oneindig voortlopende reeksen waarvan de mate van convergentie a priori niet vaststaat, een meer klassieke analytische weg gevolgd.

2. Berekening voor $n = 0$

Als inleiding tot de berekening van de reeks (7) onderzoeken wij allereerst de reeks gegeven door

$$\sum_{r=1}^{\infty} \frac{J_\nu(2\pi r z)}{(\pi r z)^\nu} \cos(2\pi r y) \tag{8}$$

waarin J_ν de Besselfunctie van de eerste soort en orde ν voorstelt en zowel ν, y als z reëel verondersteld zijn. Voor z beperken we ons tot waarden groter dan nul, terwijl op grond van de periodiciteit van de goniometrische functie zonder verlies van algemeenheid y beperkt kan blijven tot het interval $0 \leqq y \leqq \frac{1}{2}$.

Daartoe bepalen we allereerst de waarde van de uitdrukking

$$\sum_{r=1}^{R} \frac{J_\nu(2\pi r z)}{(\pi r z)^\nu} \cos(2\pi r y) \tag{9}$$

Op grond nu van de integraalbetrekking van Poisson voor de Besselfunctie van de eerste soort en orde ν:

$$J_\nu(x) = 2 \frac{\left(\frac{1}{2}x\right)^\nu}{\Pi\left(\nu - \frac{1}{2}\right)\Pi\left(-\frac{1}{2}\right)} \int_0^{\frac{1}{2}\pi} \cos\{x \sin(\phi)\} \cos^{2\nu}(\phi)\, \mathrm{d}\phi$$

geldig voor $\nu > -\frac{1}{2}$ kan voor (9) geschreven worden:

$$\sum_{r=1}^{R} \frac{J_\nu(2\pi rz)}{(\pi rz)^\nu} \cos(2\pi ry) = \frac{2}{\Pi\left(\nu-\frac{1}{2}\right)\Pi\left(-\frac{1}{2}\right)} \times$$

$$\times \int_0^{\frac{1}{2}\pi} \sum_{r=1}^{R} \cos\{2\pi rz \sin(\phi)\} \cos(2\pi ry) \cos^{2\nu}(\phi)\, d\phi \qquad (10)$$

Na sommering van de eindige reeks onder het integraalteken en onder gebruikmaking van de bekende relatie voor de Π-functies:

$$\frac{\Pi(x-1)\,\Pi(y-1)}{\Pi(x+y-1)} = 2\int_0^{\frac{1}{2}\pi} \cos^{2x-1}(\phi) \sin^{2y-1}(\phi)\, d\phi, \qquad (11)$$

geldig voor $x > 0$ en $y > 0$, ofwel $\nu > -\frac{1}{2}$, gaat (10) over in:

$$\sum_{r=1}^{R} \frac{J_\nu(2\pi rz)}{(\pi rz)^\nu} \cos(2\pi ry) + \frac{1}{2\Pi(\nu)} = \frac{1}{2\Pi\left(\nu-\frac{1}{2}\right)\Pi\left(-\frac{1}{2}\right)} \times$$

$$\left[\int_0^{\frac{1}{2}\pi} \frac{\sin\{(2R+1)\pi(z\sin(\phi)+y)\}}{\sin\{\pi(z\sin(\phi)+y)\}} \cos^{2\nu}(\phi)\, d\phi + \right.$$

$$\left. + \int_0^{\frac{1}{2}\pi} \frac{\sin\{(2R+1)\pi(z\sin(\phi)-y)\}}{\sin\{\pi(z\sin(\phi)-y)\}} \cos^{2\nu}(\phi)\, d\phi \right] \qquad (12)$$

Wijziging van de integratievariabele in de eerste integraal van (12) door $t = z\sin(\phi) + y$ te stellen en in de tweede integraal door de substitutie $t = z\sin(\phi) - y$ en overgang tot de limiet $R \to \infty$ geeft als resultaat

$$\sum_{r=1}^{\infty} \frac{J_\nu(2\pi rz)}{(\pi rz)^\nu} \cos(2\pi ry) + \frac{1}{2\Pi(\nu)} =$$

$$= \lim_{R\to\infty} \frac{1}{2z^{2\nu}\Pi\left(\nu-\frac{1}{2}\right)\Pi\left(-\frac{1}{2}\right)} \times$$

$$\left[\int_0^{z+y} \frac{\sin\{(2R+1)\pi t\}}{\sin(\pi t)} \{z^2-(t-y)^2\}^{\nu-\frac{1}{2}}\, dt + \right.$$

$$\left. + \int_0^{z-y} \frac{\sin\{(2R+1)\pi t\}}{\sin(\pi t)} \{z^2-(t+y)^2\}^{\nu-\frac{1}{2}}\, dt \right] \qquad (13)$$

Afhankelijk nu van de waarden van $(z+y)$ en $(z-y)$ kunnen drie gevallen onderscheiden worden:

a. $0 < z < y \leqq \frac{1}{2}$. In dit geval is $0 < (z+y) < 1$ en $-\frac{1}{2} < (z-y) < 0$, zodat door tekenwijziging in de tweede integraal voor de som van beide integralen geschreven kan worden:

$$\lim_{R\to\infty} \frac{1}{2z^{2\nu}\Pi\left(\nu-\frac{1}{2}\right)\Pi\left(-\frac{1}{2}\right)}$$

$$\int_{y-z}^{y+z} \frac{\sin\{(2R+1)\pi t\}}{\sin(\pi t)} \{z^2-(t-y)^2\}^{\nu-\frac{1}{2}}\, dt. \qquad (14)$$

Nu leert het onderzoek van de integraal

$$\int_{y-z}^{y+z} \frac{\{z^2-(t-y)^2\}^{\nu-\frac{1}{2}}}{\sin(\pi t)}\, dt \qquad (15)$$

dat deze na substitutie van een nieuwe variabele $\xi = (t-y+z)/(y+z-t)$ evenredig is met

$$\int_0^\infty \frac{\xi^{\nu-\frac{1}{2}}}{(\xi+1)^{2\nu+1} \sin[\pi\{y+(\xi-1)z/(\xi+1)\}]}\, d\xi. \qquad (16)$$

waarin de integrand $= 0(\xi^{-\nu-\frac{3}{2}})$ als $y \neq z$ en $= 0(\xi^{-\nu-\frac{1}{2}})$ als $y = z$. De integraaluitdrukking (16) convergeert dus absoluut als $\nu > -\frac{1}{2}$ en $y \neq z$ en als $\nu > \frac{1}{2}$ en $y = z$, zodat op grond van het theorema van Riemann-Lebesque de integraaluitdrukking (14) $= 0(1)$. Daarmede vereenvoudigt de uitdrukking (13) zich tot het eindresultaat

$$\sum_{r=1}^{\infty} \frac{J_\nu(2\pi rz)}{(\pi rz)^\nu} \cos(2\pi ry) + \frac{1}{2\Pi(\nu)} = 0 \qquad (17)$$

als $-\frac{1}{2} < \nu \leqq \frac{1}{2}$ en $0 < z < y \leqq \frac{1}{2}$ en als $\nu > \frac{1}{2}$ en $0 < z \leqq y \leqq \frac{1}{2}$.

b. $0 \leqq y < z \leqq \frac{1}{2}$. In dit geval is $0 < (z+y) < 1$ en $0 < (z-y) < \frac{1}{2}$ en beschouwing van de beide Dirichlet-integralen van de tweede soort in de uitdrukking (13) leert dat gezien het monotone verloop van de functie

$$\{z^2-(t-y)^2\}^{\nu-\frac{1}{2}} \qquad (18)$$

voor $R \to \infty$ de waarde van de integraal gelijk is aan

$$\frac{1}{2}\left\{z^2-(t-y)^2\right\}_{t=0}^{\nu-\frac{1}{2}} + \frac{1}{2}\left\{z^2-(t+y)^2\right\}_{t=0}^{\nu-\frac{1}{2}} = (z^2-y^2)^{\nu-\frac{1}{2}}$$

waarmede (13) overgaat in het eindresultaat

$$\sum_{r=1}^{\infty} \frac{J_\nu(2\pi rz)}{(\pi rz)^\nu} \cos(2\pi ry) + \frac{1}{2\Pi(\nu)} =$$

$$= \frac{\left(z^2-y^2\right)^{\nu-\frac{1}{2}}}{2z^{2\nu}\Pi\left(\nu-\frac{1}{2}\right)\Pi\left(-\frac{1}{2}\right)} \qquad (19)$$

als $\nu > -\frac{1}{2}$ en $0 \leqq y < z \leqq \frac{1}{2}$.

c. $0 \leqq y \leqq \frac{1}{2} < z$. In dit geval is $(z+y) > 1$ en $(z-y) > \frac{1}{2}$. Stel nu dat $(z+y)$ begrensd wordt door het interval $k < (z+y) < (k+1)$, waarin k een geheel en positief getal ($k = 1, 2$, enz.) en $(z-y)$ begrensd wordt door het interval $m < (z-y) < (m+1)$, waarin m eveneens een geheel en positief getal ($m = 1, 2$, enz.). Beschouwt men nu vervolgens de eerste Dirichletintegraal, dan ziet men dat de integratievariabele t alle gehele waarden van 1 tot en met k doorloopt. Splitst men nu vervolgens de integraal in de intervallen $0 \to 1$, $1 \to 2$, ..., $(k-1) \to k$, $k \to (k+1)$, dan levert iedere deelintegraal een bijdrage gelijk aan de halve som van de waarden van de functie (18) ter plaatse van de

ondergrens en de bovengrens, en het totaal wordt gegeven door de eindige reeks

$$\frac{1}{2}(z^2 - y^2)^{\nu-\frac{1}{2}} + \sum_{k=1}^{k} \{z^2 - (k-y)^2\}^{\nu-\frac{1}{2}} \tag{20}$$

Past men nu eenzelfde beschouwing toe op de tweede Dirichlet-integraal van (13) en sommeert men deze resultaten met die gevonden onder (20) dan luidt het gezochte eindresultaat:

$$\sum_{r=1}^{\infty} \frac{J_\nu(2\pi rz)}{(\pi rz)^\nu} \cos(2\pi ry) + \frac{1}{2\Pi(\nu)} =$$

$$= \frac{1}{2z^{2\nu}\Pi\left(\nu-\frac{1}{2}\right)\Pi\left(-\frac{1}{2}\right)}\left[(z^2-y^2)^{\nu-\frac{1}{2}} + \right.$$

$$\left. + \sum_{k=1}^{k} \{z^2-(k-y)^2\}^{\nu-\frac{1}{2}} + \sum_{m=1}^{m} \{z^2-(m+y)^2\}^{\nu-\frac{1}{2}}\right] \tag{21}$$

als $\nu > -\frac{1}{2}$ en k en m gehele positieve getallen, zó gekozen dat $k < (z+y) < (k+1)$ en $m < (z-y) < (m+1)$.

3. Discussie

De bovenstaande resultaten, samengevat in de vergelijkingen (17), (19) en (21) leiden tot een groot aantal interessante resultaten voor zekere waarden van z en y waarvan enkele in de literatuur bekend zijn. Zo levert (17) voor $\nu = \frac{1}{2}$ de reeks [7]:

$$\sum_{r=1}^{\infty} (-)^r \frac{J_\nu(2\pi rz)}{(\pi rz)^\nu} + \frac{1}{2\Pi(\nu)} = 0 \tag{22}$$

als $-\frac{1}{2} < \nu \leqq \frac{1}{2}$ en $0 < z < \frac{1}{2}$ en als $\nu > \frac{1}{2}$ en $0 < z \leqq \frac{1}{2}$. Hieruit vloeit als bijzonder geval voort de reeks:

$$\sum_{r=1}^{\infty} (-)^r J_0(2\pi rz) + \frac{1}{2} = 0 \tag{23}$$

geldig voor $0 < z < \frac{1}{2}$ [4].

Voor waarden van z, groter dan $\frac{1}{2}$ geldt, indien $y = \frac{1}{2}$, de betrekking $k = (m+1)$, waarmede (21) overgaat in:

$$\sum_{r=1}^{\infty} (-)^r \frac{J_\nu(2\pi rz)}{(\pi rz)^\nu} + \frac{1}{2\Pi(\nu)} = \frac{1}{z^{2\nu}\Pi\left(\nu-\frac{1}{2}\right)\Pi\left(-\frac{1}{2}\right)} \times$$

$$\times \sum_{n=1}^{n} \left\{z^2 - \left(\frac{2n-1}{2}\right)^2\right\}^{\nu-\frac{1}{2}} \tag{24}$$

voor $z > 0$ en wel zodanig dat $(n-\frac{1}{2}) < z < (n+\frac{1}{2})$, een resultaat dat eerder gevonden werd door Nielsen [6].

Stelt men daarentegen in (19) $y = 0$, dan ontstaat de reeks:

$$\sum_{r=1}^{\infty} \frac{J_\nu(2\pi rz)}{(\pi rz)^\nu} + \frac{1}{2\Pi(\nu)} = \frac{1}{2z\Pi\left(\nu-\frac{1}{2}\right)\Pi\left(-\frac{1}{2}\right)} \tag{25}$$

geldig voor $\nu > -\frac{1}{2}$ en $0 < z \leqq \frac{1}{2}$, een resultaat dat eveneens van Nielsen afkomstig is [6].

4. Berekening voor $n > 0$

Tot slot kan nu de gevraagde reeks (7) uit de berekende reeks (8) afgeleid worden door op de linker- en rechterleden van de eindresultaten (17), (19) en (21) n malen de operator

$$\frac{1}{z}\frac{\mathrm{d}}{\mathrm{d}z}\left(z\frac{\mathrm{d}}{\mathrm{d}z} + \nu\right) \tag{26}$$

toe te passen en vervolgens de macht van de noemer die door het herhaalde differentiëren tot $\nu - 2n$ gedaald is, weer gelijk aan ν te stellen. Men verkrijgt dan uit (17) de reeks

$$\sum_{r=1}^{\infty} \frac{J_{2n+\nu}(2\pi rz)}{(\pi rz)^\nu} \cos(2\pi ry) = 0 \tag{27}$$

met n geheel en > 0, geldig als $-\frac{1}{2} < \nu \leqq \frac{1}{2}$ en $0 \leqq z < y \leqq \frac{1}{2}$ en als $\nu > \frac{1}{2}$ en $0 \leqq z \leqq y \leqq \frac{1}{2}$. De reeks (19) levert vervolgens op

$$\sum_{r=1}^{\infty} \frac{J_{2n+\nu}(2\pi rz)}{(\pi rz)^\nu} \cos(2\pi ry) = \frac{(-)^n}{2^{2n+1}z\Pi\left(2n+\nu-\frac{1}{2}\right)\Pi\left(-\frac{1}{2}\right)} \times$$

$$\times \frac{1}{y^{2n}}\left(\frac{\mathrm{d}}{\mathrm{d}(z^{-1})}\right)^{2n}\left\{1-\frac{y^2}{z^2}\right\}^{2n+\nu-\frac{1}{2}} \tag{28}$$

met n geheel en > 0, geldig als $\nu > -\frac{1}{2}$ en $0 \leqq y < z \leqq \frac{1}{2}$, terwijl tot slot de reeks (21) als resultaat oplevert:

$$\sum_{r=1}^{\infty} \frac{J_{2n+\nu}(2\pi rz)}{(\pi rz)^\nu} \cos(2\pi ry) = \frac{(-)^n}{2^{2n+1}z\Pi\left(2n+\nu-\frac{1}{2}\right)\Pi\left(-\frac{1}{2}\right)} \times$$

$$\times \left[\frac{1}{y^{2n}}\left(\frac{\mathrm{d}}{\mathrm{d}(z^{-1})}\right)^{2n}\left\{1-\frac{y^2}{z^2}\right\}^{2n+\nu-\frac{1}{2}} + \right.$$

$$+ \sum_{k=1}^{k} \frac{1}{(k-y)^{2n}}\left(\frac{\mathrm{d}}{\mathrm{d}(z^{-1})}\right)^{2n}\left\{1-\frac{(k-y)^2}{z^2}\right\}^{2n+\nu-\frac{1}{2}} + \tag{29}$$

$$\left. + \sum_{m=1}^{m} \frac{1}{(m+y)^{2n}}\left(\frac{\mathrm{d}}{\mathrm{d}(z^{-1})}\right)^{2n}\left\{1-\frac{(m+y)^2}{z^2}\right\}^{2n+\nu-\frac{1}{2}}\right]$$

met n geheel en > 0, geldig als $\nu > -\frac{1}{2}$ en k en m gehele positieve getallen zo gekozen, dat $k < (z+y) < (k+1)$ en $m < (z-y) < (m+1)$. De functies in het rechterlid, hier weergegeven in operatorvorm, kunnen herleid worden tot vele bekende functies, zoals die van Gegenbauer in het meest algemene geval, polynomen van Legendre voor $\nu = \frac{1}{2}$, en polynomen van Tchebycheff voor $\nu = 1$, terwijl voorts ook asymptotische benaderingen voor grote waarden van z gevonden kunnen worden. Het zou echter te ver voeren om deze mogelijkheden hier in detail te vermelden.

5. Slotbeschouwing

Het eindresultaat, vervat in (29) kan ook – zij het in iets andere vorm – gegeven worden door niet met behulp van (26) naar de variabele z te differentiëren, doch door directe differentiatie van (17), (19) en (21) naar y. Men verkrijgt dan bijv. uit (21) als resultaat:

$$\sum_{r=1}^{\infty} \frac{J_{2n+\nu}(2\pi rz)}{(\pi rz)^{\nu}} \cos(2\pi ry) =$$

$$= \frac{(-)^n}{2^{2n+1} z^{2n+2\nu} \Pi\left(2n+\nu-\frac{1}{2}\right) \Pi\left(-\frac{1}{2}\right)} \times \left(\frac{\mathrm{d}}{\mathrm{d}y}\right)^{2n} \left[(z^2-y^2)^{2n+\nu-\frac{1}{2}} + \right.$$

$$\left. + \sum_{k=1}^{k} \{z^2-(k-y)^2\}^{2n+\nu-\frac{1}{2}} + \sum_{m=1}^{m} \{z^2-(m+y)^2\}^{2n+\nu-\frac{1}{2}} \right] \quad (30)$$

onder dezelfde condities als bij (29) vermeld.

Differentieert men niet $2n$-maal, doch $(2n-1)$-maal ($n=1, 2,$ enz.) dan ontstaan reeksen met $\phi(2\pi ry) = \sin(2\pi ry)$, waarvan door uitwerking van (17) als voorbeeld genoemd moge worden:

$$\sum_{r=1}^{\infty} \frac{J_{2n+\nu-1}(2\pi rz)}{(\pi rz)^{\nu}} \sin(2\pi ry) = 0 \quad (31)$$

met n geheel en > 0, als $-\frac{1}{2} < \nu \leqq \frac{1}{2}$ en $0 \leqq z < y \leqq \frac{1}{2}$ en als $\nu > \frac{1}{2}$ en $0 \leqq z \leqq y \leqq \frac{1}{2}$. Het is duidelijk dat ook hier weer een terugvoering naar vele bekende functies mogelijk is.

Literatuur

[1] J. P. Schouten en H. W. F. van 't Groenewout: Analysis of distortion in pulse-code-modulation systems. Applied Scientific Research, Section B, Vol. II 1952, p. 277 ff.

[2] H. W. F. van 't Groenewout: Some remarks on non-linear distortion in pcm-systems. Symposium Londen, 1953.

[3] H. W. F. van 't Groenewout: A note on Schlömilch series. Intern rapport Afdeling der Electrotechniek, Technische Hogeschool, Delft, maart 1952.

[4] G. N. Watson: A treatise on the theory of Bessel-functions. Cambridge University Press 1952, p. 634.

[5] G. N. Watson, loc. cit. p. 636.

[6] N. Nielsen: Handbuch der Zylinderfunktionen, 1904.

[7] G. N. Watson, loc. cit., p. 635.

539.31

The surface line source problem in elastodynamics



by prof. dr. ir. A. T. de Hoop, Delft University of Technology,

Department of Electrical Engineering

Summary: At the plane boundary of a semi-infinite, homogeneous, isotropic, perfectly elastic solid a pulsed line source of normal pressure generates a two-dimensional elastic wave motion in the solid. The displacement vector of this elastic wave motion is calculated at any point inside the solid.

1. Introduction

The surface line source problem in elastodynamics involves the calculation of the displacement vector of the elastic wave motion in a semi-infinite, elastic solid, generated by a pulsed line source of force applied at its plane boundary. Lamb [1], who has been the first to investigate this problem, calculated the displacement vector at the boundary. In the present paper the displacement vector at any point inside the solid is determined by employing the author's modification (de Hoop [2]) of Cagniard's technique for solving seismic pulse problems.

2. Formulation of the problem

The elastic waves under consideration are small-amplitude disturbances travelling in a semi-infinite, homogeneous, isotropic, perfectly elastic solid. The physical properties of the solid are characterized by its mass density ρ and its Lamé constants λ and μ. A right-handed Cartesian co-ordinate system x, y, z is introduced such that the elastic medium occupies the half-space $-\infty < x < \infty$, $-\infty < y < \infty$, $0 < z < \infty$. A point inside the solid or at its boundary $z = 0$ is located by either its Cartesian co-ordinates or its cylindrical co-ordinates defined through

$$x = r\sin(\theta); \qquad y = y; \qquad z = r\cos(\theta) \quad (2.1)$$

with $0 \leqq r < \infty$, $-\pi/2 \leqq \theta \leqq \pi/2$. The time co-ordinate is denoted by t.

Along the line $x=0$, $-\infty<y<\infty$, $z=0$ a pulsed force is applied to the free surface of the solid. It starts to act at the instant $t=0$; prior to this instant the medium is assumed to be at rest (Fig. 1).

Let $\vec{u} = \vec{u}(x, y, z, t)$ be the displacement vector of the elastic wave. At any interior point of the solid the stress tensor, with components τ_{xx}, $\tau_{xy}, \ldots, \tau_{zz}$, is related to the displacement vector through the (linearized) constitutive relations

$$\tau_{xx} = \lambda \operatorname{div} \vec{u} + 2\mu(\partial u_x/\partial x), \text{ etc.} \quad (2.2)$$

$$\tau_{xy} = \tau_{yx} = \mu(\partial u_x/\partial y + \partial u_y/\partial x), \text{ etc.} \quad (2.3)$$

From the (linearized) equation of motion and the constitutive relations it follows that, in the absence of body forces, the displacement vector satisfies the elastodynamic wave equation

$$v_P^2 \operatorname{grad} \operatorname{div} \vec{u} - v_S^2 \operatorname{rot} \operatorname{rot} \vec{u} - \partial^2\vec{u}/\partial t^2 = \vec{0} \quad (2.4)$$

in which

$$v_P = [(\lambda + 2\mu)/\rho]^{\frac{1}{2}} \quad (2.5)$$

is the velocity of propagation of compressional or P-waves (for which $\operatorname{rot} \vec{u} = \vec{0}$) and

$$v_S = (\mu/\rho)^{\frac{1}{2}} \quad (2.6)$$

is the velocity of propagation of shear or S-waves (for which $\operatorname{div} \vec{u} = 0$).

Now we investigate the case where the amplitude of the applied force is independent of y; then the boundary conditions as $z \downarrow 0$ are independent of y. Since the elastodynamic wave equation admits y-independent solutions the generated elastic wave motion is independent of y. Hence, $\vec{u} = \vec{u}(x, z, t)$. The y-independent solutions of (2.4) separate into two classes: in

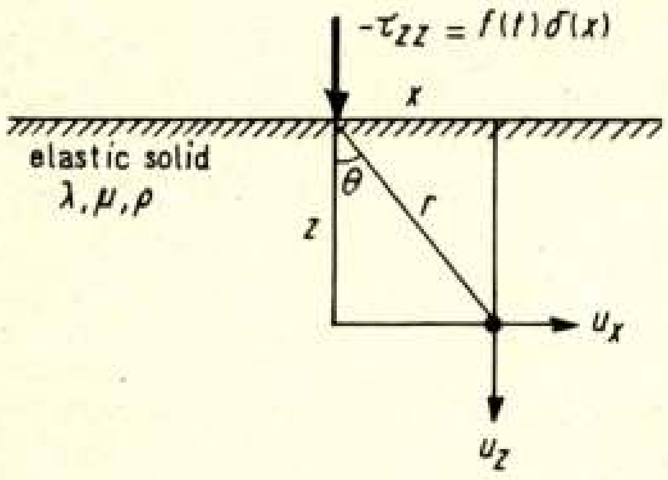


Fig. 1. The configuration to be investigated.

one of them $\vec{u}$ has only a y-component, in the other $\vec{u}$ has only an x- and a z-component. Let the plane of the boundary be 'horizontal', then the former displacement vector corresponds to an SH-wave (horizontally polarized shear wave) while the latter displacement vector corresponds to a superposition of a P-wave and an SV-wave (vertically polarized shear wave) (Bullen [3]). The calculation of the SH-wave is a simple problem in scalar wave propagation and is not considered here. The calculation of the combined P- and SV-wave motion is the subject of our further investigation.

Let $\vec{u}^{(P)}$ be the compressional part and $\vec{u}^{(SV)}$ the shear part of the relevant displacement vector, then $\vec{u} = \vec{u}^{(P)} + \vec{u}^{(SV)}$. From (2.4) it follows that $\vec{u}^{(P)}$ satisfies the two-dimensional wave equation

$$(\partial^2/\partial x^2 + \partial^2/\partial z^2 - v_P^{-2}\partial^2/\partial t^2)\vec{u}^{(P)} = \vec{0} \tag{2.7}$$

together with the auxiliary relation $\vec{i}_y \cdot \mathrm{rot}\, \vec{u}^{(P)} = 0$, i.e.

$$\partial u_x^{(P)}/\partial z - \partial u_z^{(P)}/\partial x = 0, \tag{2.8}$$

while $\vec{u}^{(SV)}$ satisfies the two-dimensional wave equation

$$(\partial^2/\partial x^2 + \partial^2/\partial z^2 - v_S^{-2}\partial^2/\partial t^2)\vec{u}^{(SV)} = \vec{0} \tag{2.9}$$

together with the auxiliary relation $\mathrm{div}\, \vec{u}^{(SV)} = 0$, i.e.

$$\partial u_x^{(SV)}/\partial x + \partial u_z^{(SV)}/\partial z = 0 \tag{2.10}$$

The calculation is further restricted to the case where the force at the boundary is applied normally, then we have

$$\lim_{z\downarrow 0} \tau_{xz} = 0 \text{ and } \lim_{z\downarrow 0} \tau_{zz} = -f(t)\delta(x) \tag{2.11}$$

where $f(t)$ denotes the amplitude of the applied pressure ($f(t) = 0$ when $-\infty < t < 0$) and $\delta(x)$ denotes the one-dimensional delta distribution. In view of the symmetry of the configuration with respect to the plane $x = 0$, u_x is an odd function of x, while u_z is an even function of x, i.e.

$$u_x(x, z, t) = -u_x(-x, z, t) \text{ and } u_z(x, z, t) = u_z(-x, z, t) \tag{2.12}$$

As a consequence, u_x and u_z need only be determined in the quarter-space $0 \leq x < \infty$, $-\infty < y < \infty$, $0 < z < \infty$. Finally, it has to be taken into account that the waves generated by the source travel away from it.

3. Method of solution

We follow the steps as indicated in [2] and subject all time-dependent quantities to a one-sided laplace transform with respect to time. For example:

$$F(s) = \int_0^\infty \exp(-st)f(t)\mathrm{d}t \tag{3.1}$$

in which s is a real, positive number large enough to ensure the convergence of integrals of the type (3.1). (It is assumed that the behaviour of $f(t)$ and $\vec{u}(x, z, t)$ as $t \to \infty$ is such that an appropriate value of s can be found.) Let $\vec{U} = \vec{U}(x, z; s)$ denote the one-sided laplace transform of $\vec{u} = \vec{u}(x, z, t)$ and let, similarly, T_{xx}, etc. be the one-sided laplace transforms of τ_{xx}, etc. Then $\vec{U}^{(P)}$ satisfies the differential equation

$$(\partial^2/\partial x^2 + \partial^2/\partial z^2 - s^2/v_P^2)\vec{U}^{(P)} = \vec{0} \tag{3.2}$$

together with the auxiliary condition

$$\partial U_x^{(P)}/\partial z - \partial U_z^{(P)}/\partial x = 0 \tag{3.3}$$

while $\vec{U}^{(SV)}$ satisfies the differential equation

$$(\partial^2/\partial x^2 + \partial^2/\partial z^2 - s^2/v_S^2)\vec{U}^{(SV)} = \vec{0} \tag{3.4}$$

together with the auxiliary condition

$$\partial U_x^{(SV)}/\partial x + \partial U_z^{(SV)}/\partial z = 0 \tag{3.5}$$

Further the constitutive relations (2.2) and (2.3) are replaced by corresponding relations involving the laplace-transformed quantities, whereas the boundary conditions (2.11) are replaced by

$$\lim_{z\downarrow 0} T_{xz} = 0 \text{ and } \lim_{z\downarrow 0} T_{zz} = -F(s)\delta(x) \tag{3.6}$$

Next we introduce the integral representations

$$\vec{U}^{(P)} = (F(s)/2\pi\mathrm{i}) \int_{-\mathrm{i}\infty}^{\mathrm{i}\infty} (p\vec{i}_x + \gamma_P\vec{i}_z)A(p) \times \exp[-s(px + \gamma_P z)]\mathrm{d}p \tag{3.7}$$

and

$$\vec{U}^{(SV)} = (F(s)/2\pi\mathrm{i}) \int_{-\mathrm{i}\infty}^{\mathrm{i}\infty} (\gamma_S\vec{i}_x - p\vec{i}_z)B(p) \times \exp[-s(px + \gamma_S z)]\mathrm{d}p \tag{3.8}$$

in which

$$\gamma_{P,S} = (1/v_{P,S}^2 - p^2)^{\frac{1}{2}} \text{ with } \mathrm{Re}(\gamma_{P,S}) \geq 0 \tag{3.9}$$

The right-hand side of (3.7) satisfies (3.2) and (3.3), the right-hand side of (3.8) satisfies (3.4) and (3.5); the choice of the square root in (3.9) will lead to waves travelling away from the origin. Substituting (3.7) and (3.8) in the boundary conditions (3.6) and using the representation

$$\delta(x) = (s/2\pi\mathrm{i}) \int_{-\mathrm{i}\infty}^{\mathrm{i}\infty} \exp(-spx)\mathrm{d}p \tag{3.10}$$

we obtain two linear, algebraic equations for $A(p)$ and $B(p)$. The solution of these equations is obtained as

$$A(p) = (1/2v_S^2 - p^2)/2\mu R(p);\ B(p) = p\gamma_P/2\mu R(p) \tag{3.11}$$

in which

$$R(p) = (1/2v_S^2 - p^2)^2 + p^2\gamma_P\gamma_S \tag{3.12}$$

At this point we observe that $A(p)$ and $B(p)$ have the following singularities in the complex p-plane: (a) branch points at $p = \pm 1/v_P$, (b) branch points at $p = \pm 1/v_S$, (c) simple poles at $p = \pm 1/v_R$ where $p = 1/v_R$ and $p = -1/v_R$ (v_R = Rayleigh wave velocity) are the simple zeros of the right-hand side of (3.12).

Now that $\vec{U}^{(P)}$ and $\vec{U}^{(SV)}$ have been determined, the transformation back to the time domain follows by changing the path of integration in (3.7) and (3.8) to a curve along which

$$\mathrm{Re}(px + \gamma_{P,S}z) = \tau > 0 \quad \text{and} \quad \mathrm{Im}(px + \gamma_{P,S}z) = 0 \tag{3.13}$$

Since this process requires the integrands to be single-valued, branch cuts are introduced in accordance with (3.9), i.e. along $\mathrm{Re}(\gamma_{P,S}) = 0$. Further, neither singularities nor branch cuts may be passed. The modified path of integration in (3.7) is then obtained as $p = \omega_P(x, z, \tau)$, together with its image $p = \omega_P^*$

with respect to the real axis, where ω_P is given by

$$\omega_P = (\tau/r)\sin(\theta) + i(\tau^2/r^2 - 1/v_P^2)^{\frac{1}{2}}\cos(\theta) \quad (3.14)$$

$$\text{with } T_P < \tau < \infty \quad \text{and} \quad T_P = r/v_P$$

The modified path of integration in (3.8) in general consists of the two parts $p = \omega_{PS}(x, z, \tau)$ and $p = \omega_S(x, z, \tau)$, together with their images $p = \omega_{PS}^*$ and $p = \omega_S^*$ with respect to the real axis, where ω_{PS} is given by

$$\omega_{PS} = (\tau/r)\sin(\theta) - (1/v_S^2 - \tau^2/r^2)^{\frac{1}{2}}\cos(\theta) + i0 \quad (3.15)$$

with $T_{PS} < \tau < T_S$ and

$$T_{PS} = T_P\sin(\theta) + T_S(1 - v_S^2/v_P^2)^{\frac{1}{2}}\cos(\theta),$$

$$T_S = r/v_S$$

and ω_S by

$$\omega_S = (\tau/r)\sin(\theta) + i(\tau^2/r^2 - 1/v_S^2)^{\frac{1}{2}}\cos(\theta) \quad (3.16)$$

$$\text{with } T_S < \tau < \infty$$

The part $p = \omega_{PS}$, together with its image $p = \omega_{PS}^*$ with respect to the real axis, is only present for points of observation in the domain $v_S/v_P < |\sin(\theta)| \leqq 1$ and is due to the presence of the branch points $p = \pm 1/v_P$ (Fig. 2). Introducing in (3.7) and (3.8) τ as variable of integration and taking into account the symmetry of the paths of integration with respect to the real axis, we obtain expressions for $\vec{U}^{(P)}$ and $\vec{U}^{(SV)}$ of the form

$$\vec{U}^{(P)} = F(s)\vec{G}^{(P)}(x, z; s) \text{ and } \vec{U}^{(SV)} = F(s)\vec{G}^{(SV)}(x, z; s) \quad (3.17)$$

where

$$\vec{G}^{(P)} = \int_{T_P}^{\infty} \exp(-s\tau)\,\vec{g}^{(P)}(x, z, \tau)\,d\tau \quad (3.18)$$

and

$$\vec{G}^{(SV)} = \int_{T_{PS}}^{T_S} \exp(-s\tau)\,\vec{g}^{(PS)}(x, z, \tau)\,d\tau + \int_{T_S}^{\infty} \exp(-s\tau)\,\vec{g}^{(S)}(x, z, \tau)\,d\tau \quad (3.19)$$

with

$$\vec{g}^{(P)} = (1/\pi)\,\mathrm{Im}\,[(\omega_P\vec{i}_x + \gamma_P(\omega_P)\vec{i}_z)\,A(\omega_P)\,(\partial\omega_P/\partial\tau)] \quad (3.20)$$

$$\vec{g}^{(PS)} = (1/\pi)\,\mathrm{Im}\,[(\gamma_S(\omega_{PS})\vec{i}_x - \omega_{PS}\vec{i}_z)\,B(\omega_{PS})\,(\partial\omega_{PS}/\partial\tau)] \quad (3.21)$$

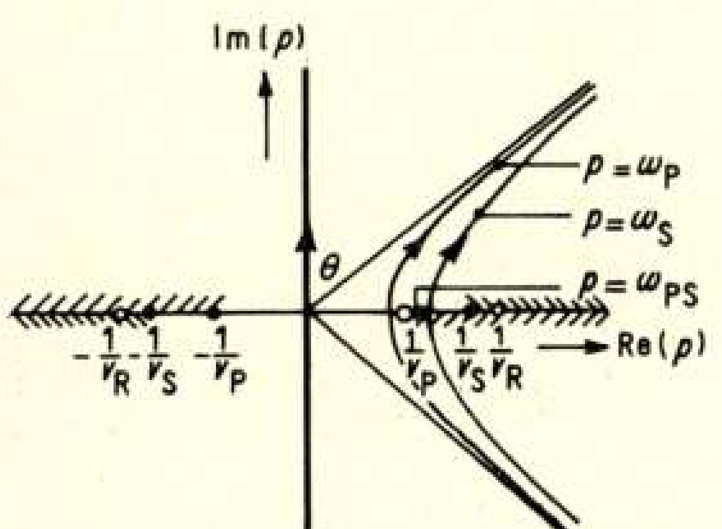


Fig. 2. The modified paths of integration in the complex p-plane.

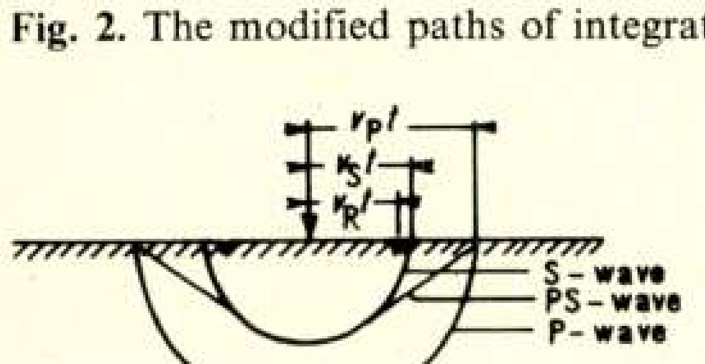


Fig. 3. Wave fronts of the P-wave, the PS-wave and the S-wave.

$$\vec{g}^{(S)} = (1/\pi)\,\mathrm{Im}\,[(\gamma_S(\omega_S)\vec{i}_x - \omega_S\vec{i}_z)\,B(\omega_S)\,(\partial\omega_S/\partial\tau)] \quad (3.22)$$

Expressions for $\vec{u}^{(P)}(x, z, t)$ and $\vec{u}^{(SV)}(x, z, t)$ are now obtained by applying the convolution theorem to (3.17) and using the uniqueness of the inverse laplace transform. The results are presented in Table 1 and Table 2.

Table 1. P-wave as a function of time.

t	$\vec{u}^{(P)}(x, z, t)$
$-\infty < t < T_P$	$\vec{0}$
$T_P < t < \infty$	$\int_{T_P}^{t} f(t-\tau)\,\vec{g}^{(P)}(x, z, \tau)\,d\tau$

Table 2. SV-wave as a function of time (the term $\vec{g}^{(PS)}$ is only present in the domain $v_S/v_P < |\sin(\theta)| \leqq 1$).

t	$\vec{u}^{(SV)}(x, z, t)$
$-\infty < t < T_{PS}$	$\vec{0}$
$T_{PS} < t < T_S$	$\int_{T_{PS}}^{t} f(t-\tau)\,\vec{g}^{(PS)}(x, z, \tau)\,d\tau$
$T_S < t < \infty$	$\int_{T_{PS}}^{T_S} f(t-\tau)\,\vec{g}^{(PS)}(x, z, \tau)\,d\tau + \int_{T_S}^{t} f(t-\tau)\,\vec{g}^{(S)}(x, z, \tau)\,d\tau$

4. Discussion of the results

The total elastodynamic wave motion $\vec{u}$ generated by the line source is given by $\vec{u} = \vec{u}^{(P)} + \vec{u}^{(SV)}$, where $\vec{u}^{(P)}$ follows from Table 1 and $\vec{u}^{(SV)}$ follows from Table 2 (see Fig. 3). The P-wave is a cylindrical wave arriving at $t = T_P$, where T_P is the travel time for P-waves travelling along a straight line from the line source to the point of observation. The SV-wave consists of a PS-conversion wave arriving at $t = T_{PS}$ and a cylindrical SV-wave arriving at $t = T_S$. The PS-conversion wave is only present in the domain $v_S/v_P < |\sin(\theta)| \leqq 1$ and precedes the cylindrical SV-wave. T_{PS} is the travel time for SV-waves travelling from the line source a distance $x - z\,(v_S/v_P)\,(1 - v_S^2/v_P^2)^{-\frac{1}{2}}$ along the boundary with P-wave velocity and then along a straight line to the point of observation with S-wave velocity; T_S is the travel time for S-waves travelling along a straight line from the line source to the point of observation. At the boundary of the solid there is a singularity in the displacement vector travelling with velocity v_R; this singularity is due to the presence of the pole at $p = \pm 1/v_R$.

The features that are characteristic for the configuration occur in $\vec{g}^{(P)}$, $\vec{g}^{(PS)}$ and $\vec{g}^{(S)}$. The numerical computation of these quantities involves only algebraic operations of the type complex.

References

[1] Lamb, H.: On the propagation of tremors over the surface of an elastic solid. Phil. Trans. Roy. Soc. (London) **A 203** (1904) 1-42.

[2] De Hoop, A. T.: A modification of Cagniard's method for solving seismic pulse problems. Appl. Sci. Res. **B 8** (1960) 349-356.

[3] Bullen, K. E.: An introduction to the theory of seismology. Cambridge University Press. 3rd edition, 1963, p. 75.

621.313.32.016

Transiënte verschijnselen bij synchrone machines

door prof. dr. ir. F. A. W. van den Burg, Afdeling der Elektrotechniek, T.H. Delft



Summary: *On transients of the synchronous machine.*
Frequently, transients of the synchronous machine have been studied by analytical methods and pictured almost exclusively by the phase portraits of the load angle.

A different portraiture is introduced with attractive features as to the insight in the dynamical behaviour due to impact loads.

Besides, the validity of the second order differential equation, which is called 'the differential equation of the synchronous machine', is discussed.

1. Inleiding

De afbeeldingen die het stationaire gedrag van elektrische machines beschrijven behoren tot de standaardafbeeldingen van de vakliteratuur. Tot zulk soort afbeeldingen kunnen voor de synchrone machines de koppel-lasthoekkarakteristieken en de stroomdiagrammen worden gerekend.

Stelt men voor de synchrone machine, onder niet te essentieel beperkende voorwaarden, het volledige stelsel op van differentiaalvergelijkingen dat het elektrodynamische gedrag van de machine geheel beschrijft, dan vindt men een stelsel van de vierde orde voor het geval dat er geen demperkooi aanwezig is en een van de zesde orde indien er wel een demperkooi aanwezig is. Na een passende assentransformatie kan men deze stelsels als autonome stelsels schrijven. Dan kunnen ook de stationaire oplossingen van de oorspronkelijke stelsels van de vierde of de zesde orde worden gevonden en de reeds genoemde karakteristieken en diagrammen worden geconstrueerd.

Voor het verkrijgen van afbeeldingen van overgangsverschijnselen die optreden na plotselinge veranderingen die op de machine worden uitgeoefend, kan men in principe twee geheel verschillende wegen inslaan. Men kan de genoemde stelsels direct bewerken met behulp van rekentuigen of men kan deze stelsels zodanig analytisch bewerken dat enig inzicht op grond van de verkregen mathematische vereenvoudigingen mogelijk wordt.

Tussen de beide uiterste wegen ligt de middenweg van de gedoseerde analytische voorbewerking van het probleem. Voor een daarop volgende bewerking met een digitaal rekentuig betekent zo'n voorbewerking veelal het traceren van de sporen waarlangs de programmering zich voor een economisch gebruik van dat hulpmiddel moet bewegen. Zo'n middenweg leidt bij het gebruik van analogonrekentuigen bijna steeds tot aanzienlijke complicaties voor de daarbij vereiste programmering.

De eerst in de afgelopen decennia ingezette snelle ontwikkeling van beide rekentechnische hulpmiddelen maakt het begrijpelijk dat vóór 1940 vooral de analytische bewerking tot op het niveau van een mathematisch interpreteerbare formulering de volle aandacht had. Reeds in 1911 leidde Dreyfus [1] een benaderende differentiaalvergelijking van de tweede orde af voor de (tweefasige) synchrone machine. Aan zijn afleiding werden vele details toegevoegd, bijvoorbeeld door Bödefeld [2] Edgerton en Zak [3], Edgerton en Fourmarier [4], Lyon en Edgerton, [5], Timascheff [6]. Hun afleidingen berusten op gevoel voor verhoudingen, fysisch inzicht, intuïtie, of hoe men dit ook noemen wil, maar zeker op een grondige praktische ervaring met deze machines. Toch moet worden gesteld dat de voorwaarden waaronder de afgeleide differentiaalvergelijkingen met een te schatten nauwkeurigheid geldig zijn, volledig in het duister bleven. Mede daarom werd reeds in 1935 het gehele stelsel differentiaalvergelijkingen met behulp van een analogonrekentuig bewerkt [7].

Een goede revisie van de mathematische formuleringen van de genoemde stelsels brachten, enigszins onafhankelijk van elkaar, Laible [8] en Concordia [9]. Op mathematisch gebied trok daarnaast de toepassing van de theorieën van de niet-lineaire differentiaalvergelijkingen op de (door Dreyfus afgeleide) vergelijking van de synchrone machine via Vlasov [10], Minorsky [11] en Stoker [12] zo sterk de aandacht dat het soms tot nieuwe mathematische beschouwingen aanleiding gaf [13], maar zeker bewerkte dat de afbeeldingen van de overgangsverschijnselen die bij de synchrone machine optreden, vooral door middel van een faseportret van de lasthoek werden weergegeven.

Dat andere afbeeldingen vaak een beter inzicht in het gedrag van overgangsverschijnselen geven zal worden geïllustreerd door, voor een viertal stootveranderingen op de synchrone machine, naast de gebruikelijke faseportretten van de lasthoek, een ongebruikelijk type van afbeelding van hetzelfde viertal veranderingen weer te geven.

Deze afbeeldingen zijn met behulp van een analogonrekentuig verkregen. Voor de met het gebruik van dat hulpmiddel weinig vertrouwde lezer zij gewezen op de 'Vakantieleergang Elektrotechniek 1959' [14] en op de recente dissertatie van Köppel [15].

Tegen de bovenstaande achtergrond zal eveneens, uitgaande van een formulering overeenkomstig aan die van Laible en Concordia, nog een analytische voorbewerking van het beschrijvende stelsel van differentiaalvergelijkingen worden behandeld waarmede de door Dreyfus reeds afgeleide benaderende beschrijving direct wordt gevonden.

2. Vergelijkingen

Voor de synchrone machine met gladde rotor zonder demperkooi (zie fig. 1) waarvan de stator is aangesloten op een star symmetrisch driefasennet vindt men als beschrijvende vergelijkingen in een genormeerd *dq*-assenstelsel dat op details na overeenkomt met het door Laible [8] gebruikte assenstelsel,

$$E = PJ + (1-\dot{\delta})\,G\Phi + \dot{\Phi};\ \Phi = KJ;\ t_{as} = j_p\ddot{\delta} + J'G\Phi \qquad (1)$$

$$X = [x_d, x_q, x_f]';\ (X = E, J, \Phi;\ x = e, j, \varphi) \qquad (2)$$

$$E = [-\sin(\delta),\ +\cos(\delta),\ +e_f]' \qquad (3)$$

$$P = \begin{bmatrix} \alpha & . & . \\ . & \alpha & . \\ . & . & \alpha_f \end{bmatrix};\ G = \begin{bmatrix} . & -1 & . \\ +1 & . & . \\ . & . & . \end{bmatrix};\ K = \begin{bmatrix} 1 & . & k \\ . & 1 & . \\ k & . & 1 \end{bmatrix} \qquad (4)$$

$$\alpha = r/X,\ \alpha_f = r_f/X_f \qquad (5)$$

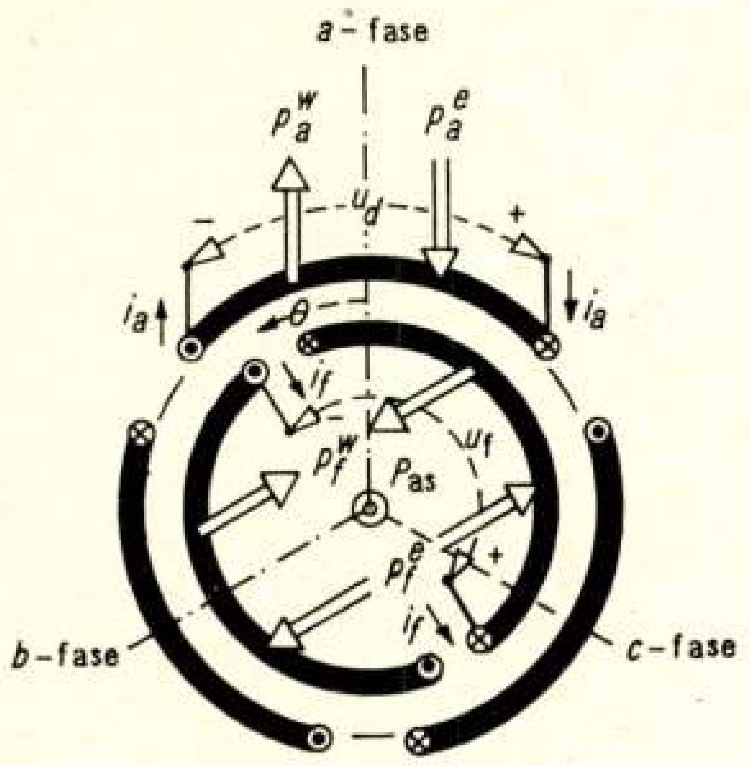


Fig. 1. Synchrone machine (motor); tekenafspraken.

r = statorweerstand (6)

X = synchrone reactantie (7)

r_f = weerstand veldwikkeling (8)

X_f = reactantie veldwikkeling bij netfrequentie $\omega_0/2\pi$ (9)

δ = lasthoek, $\dot{\delta} = \mathrm{d}\delta/\mathrm{d}\tau$, $\tau = \omega_0 t$, t = tijd in secunden (10)

k = koppelingscoëfficiënt, $\sigma = 1 - k^2$ = lekcoëfficiënt (11)

E, resp. J, resp. Φ = getransformeerde en genormeerde excitatiespanningsmatrix, resp. stroom-, resp. fluxmatrix, (12)

t_{as} = genormeerd asmoment (13)

j_p = genormeerd polair traagheidsmoment van de rotor (14)

In (1) en (2) betekent X' de matrix die uit X door transpositie ontstaat. Voorts zullen nog worden gebruikt de met α en α_f samenhangende hoeken $\bar{\alpha}$ en $\bar{\alpha}_f$ volgens,

$$\cos(\bar{\alpha}) = X/Z, \sin(\bar{\alpha}) = r/Z, Z = (r^2 + X^2)^{\frac{1}{2}}, \alpha = \tan(\bar{\alpha}) \qquad (15)$$

$$\cos(\bar{\alpha}_f) = X_f/Z_f, \sin(\bar{\alpha}_f) = r_f/Z_f, Z_f = (r_f^2 + X_f^2)^{\frac{1}{2}}$$
$$\alpha_f = \tan(\bar{\alpha}_f) \qquad (16)$$

Enig idee van de orde van grootte van de gebruikte constanten geeft tabel 1.

Tabel 1. Gebied van de machineconstanten σ, α, α_f, j_p.

Machine	σ	α	α_f	j_p
Turbo-	0,125	0,020	0,013	1500
	...	...	...	
generator	0,150	0,030	0,035	25000
Synchr.	0,150	0,010	0,035	150
	...	...	...	...
motor	0,300	0,030	0,150	500

De vergelijkingen voor een machine met uitgebouwde polen en met een volledige demperwikkeling verschillen niet principieel van (1). Zo moet men dan bijvoorbeeld in de plaats van (3) en (4) stellen:

$$P = \begin{bmatrix} \alpha_d & \cdot & \cdot & \cdot & \cdot \\ \cdot & \alpha_q & \cdot & \cdot & \cdot \\ \cdot & \cdot & \alpha_D & \cdot & \cdot \\ \cdot & \cdot & \cdot & \alpha_Q & \cdot \\ \cdot & \cdot & \cdot & \cdot & \alpha_f \end{bmatrix}; \quad G = \begin{bmatrix} \cdot & -1 & \cdot & \cdot & \cdot \\ +1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} \qquad (17)$$

$$K = \begin{bmatrix} 1 & \cdot & k_{dD} & \cdot & k_{df} \\ \cdot & 1 & \cdot & k_{qQ} & \cdot \\ k_{dD} & \cdot & 1 & \cdot & k_{Df} \\ \cdot & k_{qQ} & \cdot & 1 & \cdot \\ k_{df} & \cdot & k_{Df} & \cdot & 1 \end{bmatrix}$$

Voor het verdere betoog zijn deze verschillen niet van wezenlijk belang.

3. Stroomdiagram

Het stelsel (1) laat een particuliere oplossing toe met constante $x = x^c$, ($x = e, j, \varphi, \delta$), die samenhangen volgens

$$E^c = PJ^c + G\Phi^c; \quad \Phi^c = KJ^c; \quad t_{\mathrm{as}}^c = J^{c'} G \Phi^c = kj_f^c j_q^c \qquad (18)$$

Met (17) kan men voor deze stationaire toestand allerhande van technisch belang zijnde grootheden uitdrukken in δ^c, de constante lasthoek en in j_f^c, de constante bekrachtigingsstroom, waarbij in deze genormeerde grootheden de combinatie kj_f^c voorstelt dat deel van de magnetomotorische kracht van de veldwikkeling dat maximaal met een d- of q-statorfase kan zijn gekoppeld.

De onderlinge samenhang van de bedoelde grootheden wordt weergegeven in het bekende cirkeldiagram van de synchrone machine (zie fig. 2). Dit diagram zal tevens worden gebruikt voor het afbeelden van overgangsverschijnselen. Daarom worden nog gedefinieerd:

$$e_s = (e_d^2 + e_q^2)^{\frac{1}{2}}, (\text{n.b.}: e_s = e_s^c = 1); j_s = (j_d^2 + j_q^2)^{\frac{1}{2}} \qquad (19)$$

$$p_s^e = e_d j_d + e_q j_q = -j_d \sin(\delta) + j_q \cos(\delta) = j_s \cos(\varphi) \qquad (20)$$

$$p_s^{bl} = e_d j_q - e_q j_d = -j_d \cos(\delta) - j_q \sin(\delta) = j_s \sin(\varphi) \qquad (21)$$

waarmee, juist voor niet-stationaire toestanden, het aan de stator toegevoerde wattvermogen p_s^e en het blindvermogen p_s^{bl} als functie van de tijd τ zijn gedefinieerd en ook de fasehoek $\varphi = \varphi(\tau)$ is vastgelegd. Deze definities geven de beschikking over paren van variabele grootheden die in de stationaire toestand als constante grootheden bij elkaar behoren en van technisch belang zijn.

Het diagram van fig. 2 verschilt slechts van het gebruikelijke stroomdiagram door het gebruik van genormeerde grootheden en door een draaiing van het gebruikelijke diagram over $\pm \pi/2$ en wel + of − afhankelijk van het gekozen standaardwerk waaruit men zo'n stroomdiagram ontleend kan denken. Ook het aflezen van het (constante) asvermogen p_{as}^c en van de in de stator gedissipeerde warmtestroom $p_s^{w,c}$ komt in fig. 2 op een niet gebruikelijke manier tot stand. In de gekozen stand kan men tenslotte, door de aard van de normering, het diagram opvatten als een stroom-, een vermogens- of als een admittantiediagram.

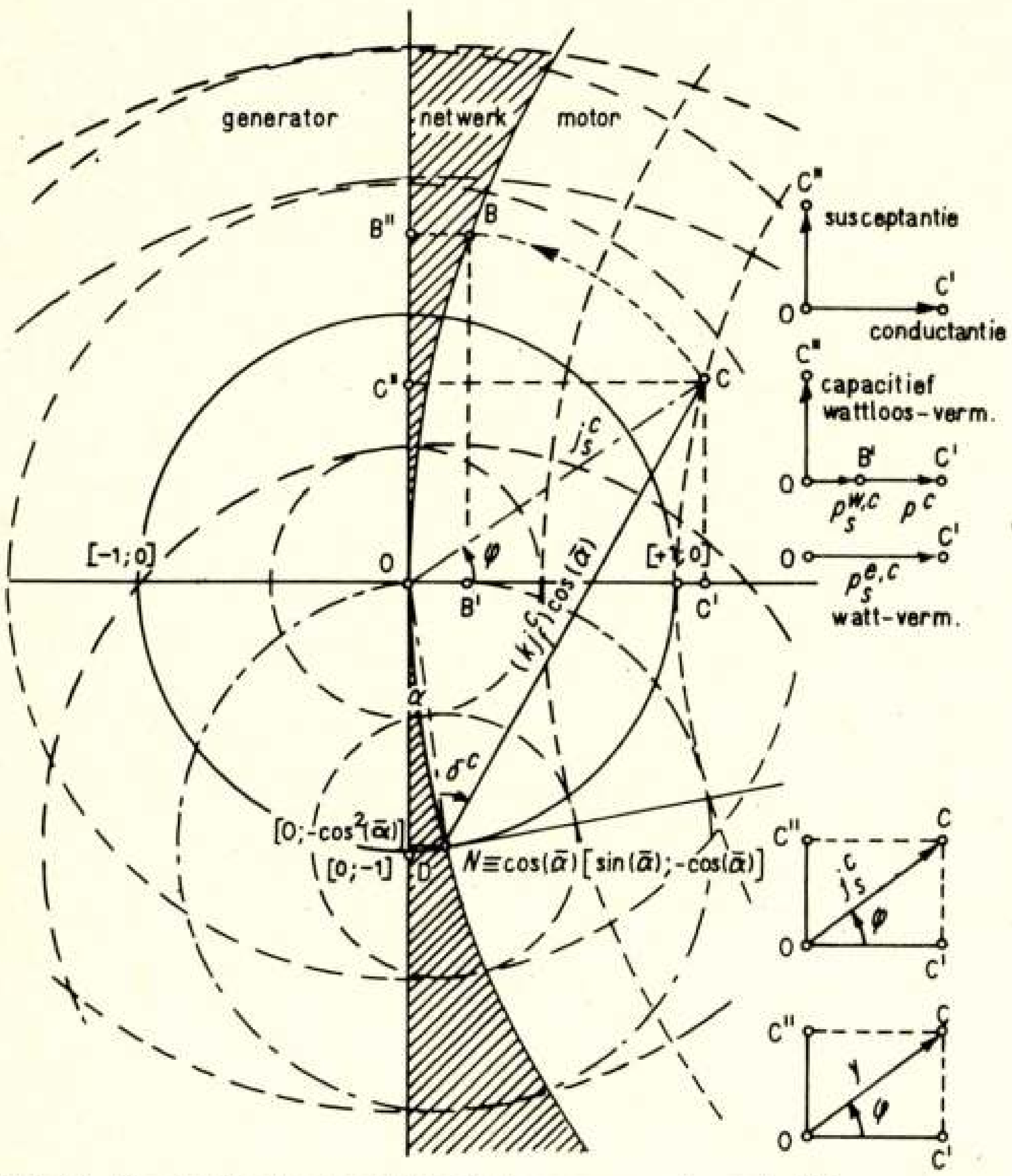


Fig. 2. De synchrone machine; het genormeerde cirkeldiagram voor de stationaire toestanden.

4. Afbeelding van stootveranderingen

Uit (1) tot en met (4) volgen voor

$$-\dot{\varphi}_d = \alpha j_d - (1-\dot{\delta})\,\varphi_q + \sin(\delta) \tag{22}$$

$$-\dot{\varphi}_q = \alpha j_q + (1-\dot{\delta})\,\varphi_d - \cos(\delta) \tag{23}$$

$$-\dot{\varphi}_f = \alpha_f j_f - e_f{}^c = \alpha_f\,(j_f - j_f{}^c) \tag{24}$$

$$\varphi_d = j_d + kj_f,\ \varphi_q = j_q,\ \varphi_f = kj_d + j_f \tag{25}$$

$$-j_p\,\ddot{\delta} = kj_f\,j_q - t_{as}^c \tag{26}$$

de vergelijkingen voor de machine met een constante bekrachtiging en een constant lastmoment. Het stelsel (19) tot en met (26) werd, voor een 3-kVA motor met

$$\alpha = 0{,}03;\ \alpha_f = 0{,}035;\ \sigma = 0{,}30;\ j_p = 250 \tag{27}$$

met behulp van een analogonrekentuig gesimuleerd.

Uitgaande van de begintoestand ($p_{as}^c = t_{as}{}^c = 1{,}44$; $kj_f{}^c\cos(\alpha) = 2$) werden de dynamische karakteristieken geregistreerd die het verband aangeven tussen enerzijds $p_s{}^{bl}(\tau)$ en $p_s{}^e(\tau)$ (of, met andere woorden, tussen $j_s(\tau)$ en $\varphi(\tau)$, zie fig. 3) en anderzijds tussen $kj_f(\tau)\cos(\bar\alpha)$ en $\delta(\tau)$ (zie fig. 4), en wel na het aanbrengen van de laststoten (geval a en b), alsook na het aanbrengen van de bekrachtigingsstoten (geval c en d).

Omdat het stelsel (19) tot en met (26) een differentiaalvergelijking van de vierde orde omvat, kan niet worden verwacht dat de overeenkomstige dynamische karakteristieken van de fig. 3 en 4 elkaar dekken. Toch blijkt dit in hoge mate het geval te zijn. Dit betekent dat het genoemde stelsel redelijk moet kunnen worden benaderd door één (niet-lineaire) differentiaalvergelijking van de tweede orde.

In fig. 5 zijn de bij de gevallen a, b, c en d behorende faseportretten van de lasthoek δ weergegeven. De onderscheiding van het subtransiënte gebied – waarin $\dot\delta$ bij benadering constant blijft – en het transiënte gebied is voor geval c en d uit de fig. 3 en 4 nog mogelijk, maar niet in de fig. 5. De afbeeldingen die zijn gesuperponeerd over het cirkeldiagram geven dan ook meer informatie en een beter inzicht in het gedrag van de overgangsverschijnselen.

5. Benadering

Uit (1) volgt voor E:

$$E = (P+GK)\,J - [(\dot{\delta}G - \varepsilon\frac{d}{d\tau})\,K]\,J;\ \varepsilon = \mathrm{diag}\,[1,1,1] \tag{28}$$

De schrijfwijze van (28) is zodanig ingericht dat voor de stationaire toestand met $\delta = \delta^c$ en (o.a.) $J = J^c$ de tweede term uit het rechter lid van (28) verdwijnt.
Men noemt $J^{(o)}$ uit

$$E = (P + GK)\,J^{(o)} \tag{29}$$

de quasistationaire beschrijving van de spanningsvergelijking (28) en het daarmede bepaalde moment

$$t_{as} = j_p\ddot{\delta} + kj_f\cos(\bar\alpha)\,[\sin(\delta + \bar\alpha) - kj_f\sin(\bar\alpha)] \tag{30}$$

de quasistationaire momentvergelijking. Als $e_f = e_f{}^c$ en $t_{as} = t_{as}{}^c$ kan men (30) vervangen door

$$j_p\ddot{\delta} + kj_f{}^c\cos(\bar\alpha)\,[\sin(\delta+\bar\alpha) - \sin(\delta^c + \bar\alpha)] = 0 \tag{31}$$

een vergelijking die de ongedempte beweging van de machine beschrijft en aanleiding geeft tot faseportretten van $\delta + \bar\alpha$ die door Stoker [12] uitvoerig zijn behandeld.

Men kan daarom verwachten dat als men voor J in de tweede term van (28) de $J^{(o)}$ uit (29) substitueert, de dan te bepalen stroommatrix uit (28) een dempingsterm in de momentvergelijking introduceert.
Stelt men nu

$$A = (P+GK)^{-1};\qquad D = (\dot{\delta}G - \varepsilon\frac{d}{d\tau})\ K \tag{32}$$

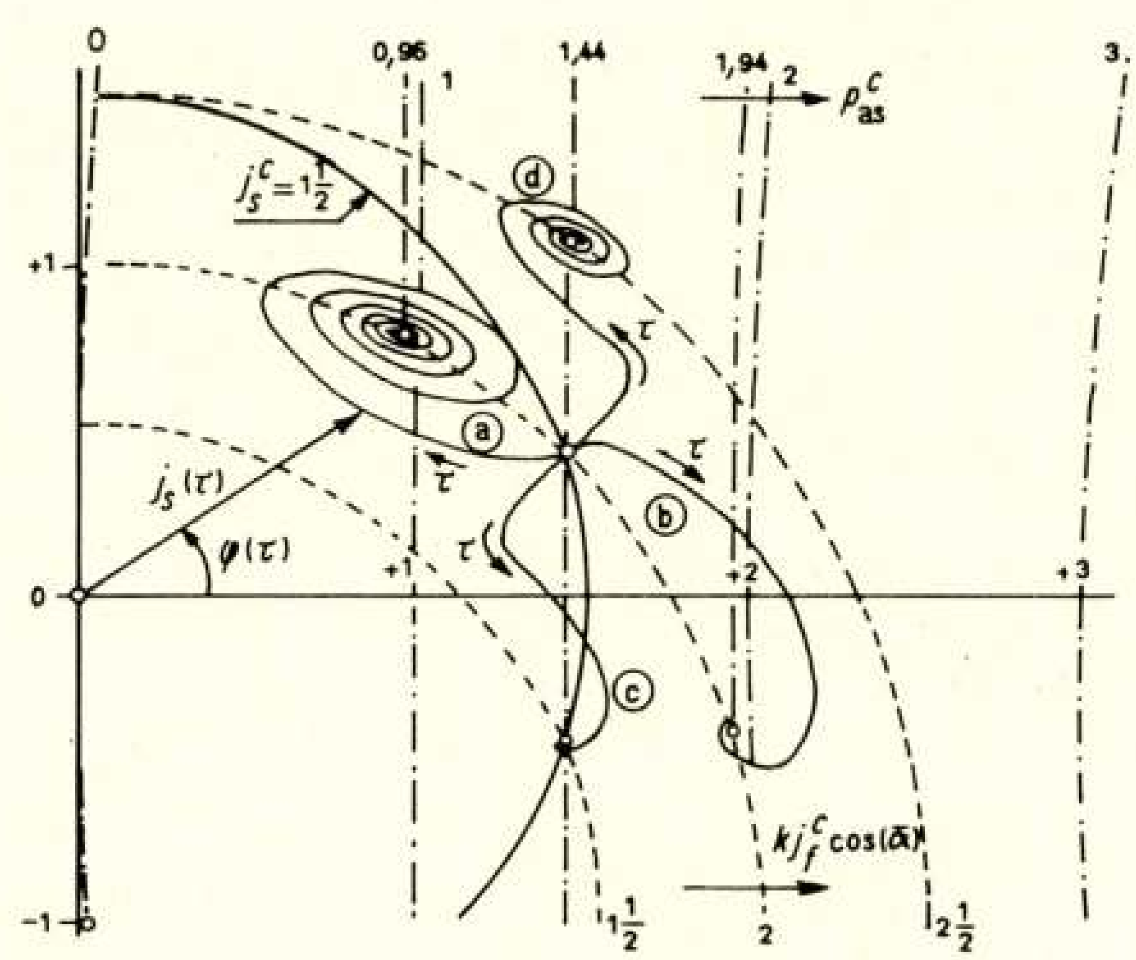


Fig. 3. Dynamische karakteristieken van ($p_s{}^{bl}$; p_s^e), dus ook van (j_s; φ), uitgaande van [$p_{as}^c = 1{,}44$; $kj_f^c\cos(\bar\alpha) = 2$], en wel:

a. met sprong van p_{as}^c van 1,44 naar 0,96;
b. met sprong van p_{as}^c van 1,44 naar 1,94;
c. met sprong van $kj_f^c\cos(\bar\alpha)$ van 2 naar $1\frac{1}{2}$;
d. met sprong van $kj_f^c\cos(\bar\alpha)$ van 2 naar $2\frac{1}{2}$.

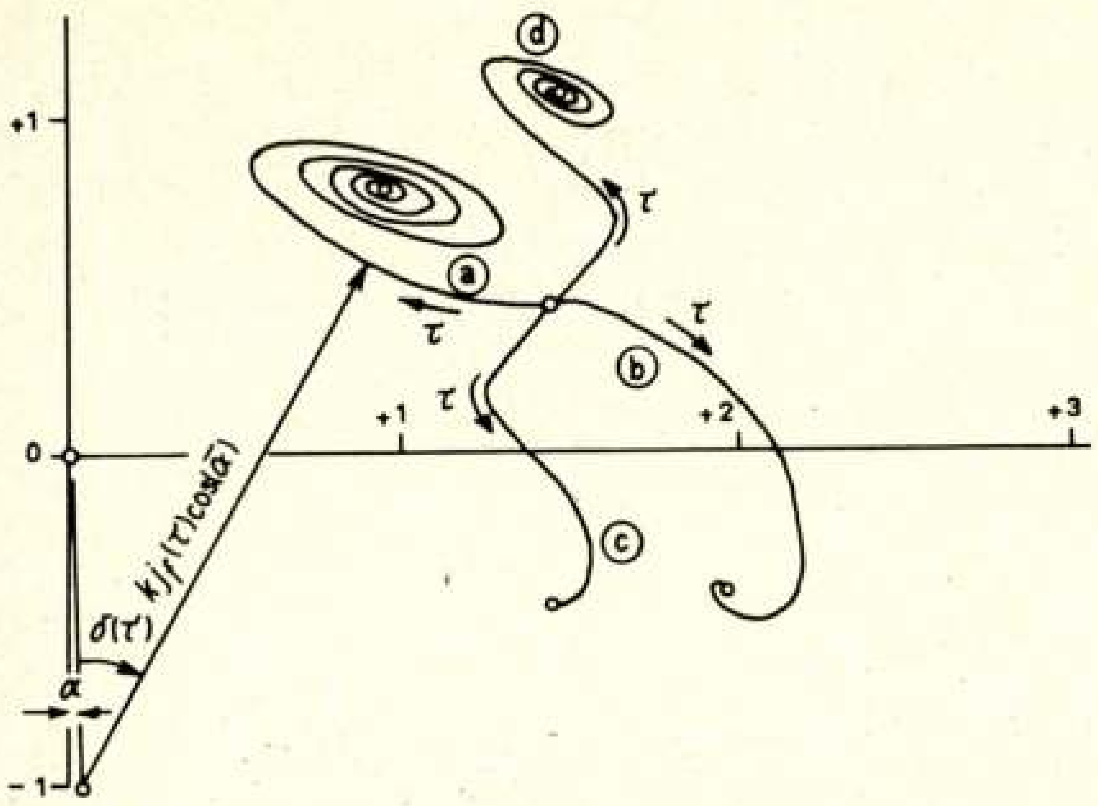


Fig. 4. Dynamische karakteristieken van ($kj_f \cos(\bar{\alpha})$; δ), uitgaande van [$p_{as}^c = 1{,}44$; $kj_f^c \cos(\bar{\alpha}) = 2$], en wel de gevallen a, b, c en d (zie fig. 3).

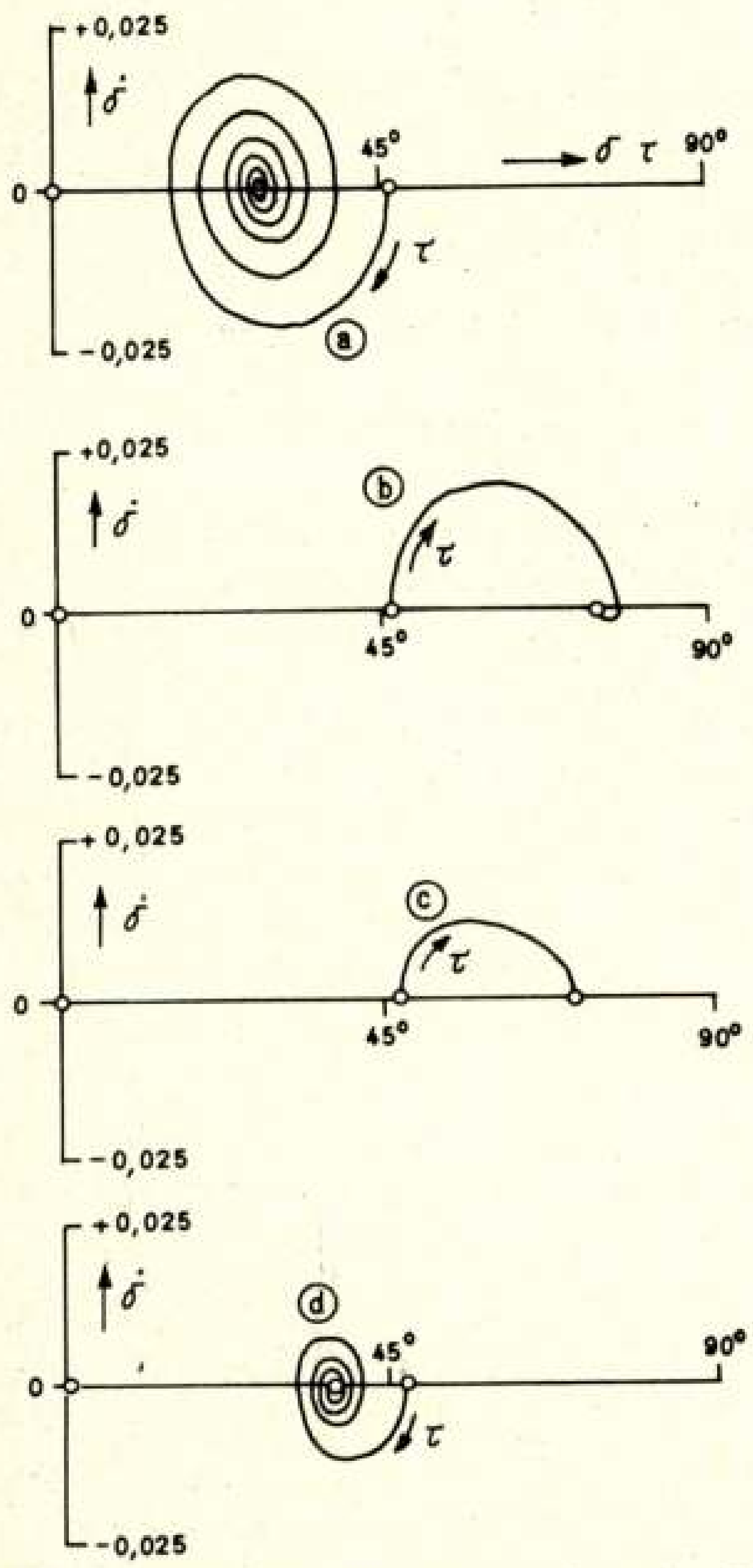


Fig. 5. Fasediagrammen voor de onder fig. 3 en 4 vermelde gevallen a, b, c en d.

dan is A een constante matrix en D een matrixoperator en kan men (28) schrijven als

$$J = AE + ADJ \tag{33}$$

Door herhaalde substitutie van het rechterlid van (33) voor J uit dat rechterlid vindt men zo de ontwikkeling

$$J = \sum_{n=0}^{\infty} (AD)^{(n)} AE \tag{34}$$

waarvan de partiële som

$$J^{(N)} = \sum_{n=0}^{N} (AD)^{(n)} AE \tag{35}$$

voor $N = 0$ de quasistationaire $J^{(o)}$ uit (29) levert. Voor $N=1$ vindt men tenslotte als benadering van de momentvergelijking:

$$j_p \ddot{\delta} + \left\{ \frac{1}{\alpha_f} k^2 \cos^2(\bar{\alpha})[\sin^2(\delta+\bar{\alpha}) - 2kj_f^c \cdot \alpha \cdot \sin(\delta+\bar{\alpha})] - (kj_f^c)^2 \frac{\sin(4\bar{\alpha})}{4} \right\} \dot{\delta} + kj_f^c \cos(\bar{\alpha})[\sin(\delta+\bar{\alpha}) - \sin(\delta^c+\bar{\alpha})] = 0 \tag{36}$$

een vergelijking die ook door Bödefeld [2] is afgeleid en waarvan de faseportretten van $\delta + \bar{\alpha}$ in Minorsky en Stoker uitvoerig worden behandeld.

6. Een bovengrens voor $|\dot{\delta}|$ bij het toepassen van (36)

Met de notaties

$$\dot{\delta} = \omega; \quad \dot{x} = \omega\,(\mathrm{d}x/\mathrm{d}\delta) = \omega x'; \quad (x = \delta, \dot{\delta}, \ddot{\delta}, \ldots) \tag{37}$$

kan men (34) met behulp van (32) en (37) schrijven als,

$$J = \sum_{n=0}^{\infty} \left[\omega\left(AGK - AK\frac{\mathrm{d}}{\mathrm{d}\delta}\right)\right]^{(n)} AE \tag{38}$$

Men kan nu, de methode uit Bellman [16] volgende, elke componentenvergelijking van (38) majoreren door één scalaire reeks; daarbij wordt gebruik gemaakt van het begrip 'norm' van een matrix B, genoteerd als $\|B\|$. Gebruikt zal worden de euclidische vorm van zo'n matrix met als definitie

$$\|B\| = [\text{som van de kwadraten van alle elementen van } B]^{\frac{1}{2}} \tag{39}$$

Bepaald werden nu

$$\|A\| = \frac{1}{\alpha_f}\left[1 + (k^2 + 2\alpha_f^2)\cos^2(\bar{\alpha})\right]^{\frac{1}{2}} = a_1 \approx 2^{\frac{1}{2}} \alpha_f^{-1}, \tag{40}$$

$$\|AK\| = \frac{1}{\alpha_f}(1+k^2)^{\frac{1}{2}}\left[1 + \left(k^2 + \frac{2+k^2}{1+k^2}\alpha_f^2\right)\cos^2(\alpha)\right]^{\frac{1}{2}} = a_2 \approx 2\alpha_f^{-1} \tag{41}$$

$$\|AGK\| = \cos(\bar{\alpha})\left[2 + k^2\cos^2(\bar{\alpha})\right]^{\frac{1}{2}} = a_3 \approx 3^{\frac{1}{2}} \tag{42}$$

waarbij voor de benaderingen van a, a_2 en a_3 de gegevens uit tabel 1 werden gebruikt. Dan kan men volgens de genoemde methode naar aanleiding van (38) stellen

$$|j_g| \leq a_1 \sum_{n=0}^{\infty}\left[|\omega|(a_3+a_2)\left|\frac{\mathrm{d}}{\mathrm{d}\delta}\right|\right]^{(n)} \|E\|, \quad (g = d, q, f) \tag{43}$$

waarbij de absoluuttekens om $\mathrm{d}/\mathrm{d}\delta$ aangeven dat bij de differentie naar δ van E en van ω alleen de absolute waarden van de resultaten tellen.

In het bijzonder moet men in (43) lezen

$$\left|\frac{\mathrm{d}}{\mathrm{d}\delta}\right|^{(m)} \|E\| = \left[\left(\frac{\mathrm{d}^m e_d}{\mathrm{d}\delta^m}\right)^2 + \left(\frac{\mathrm{d}^m e_q}{\mathrm{d}\delta^m}\right)^2 + \left(\frac{\mathrm{d}^m e_f}{\mathrm{d}\delta^m}\right)^2\right]^{\frac{1}{2}}, \quad (m \neq 0) \tag{44}$$

$$\left|\frac{\mathrm{d}}{\mathrm{d}\delta}\right|^{(0)} \|E\| = \left[e_d^2 + e_q^2 + e_f^2\right]^{\frac{1}{2}} \tag{45}$$

hetgeen voor $E = [-\sin(\delta), +\cos(\delta), +e_f^c]'$ geeft:

$$|\frac{d}{d\delta}|^{(m)} \|E\| = 1 \text{ voor } m \neq 0 \text{ en} = \left[1+(e_f^c)^2\right]^{\frac{1}{2}} \qquad (\text{voor } m = 0) \tag{46}$$

Opdat het rechterlid van (44) beperkt blijve en daarmede de schatting volgens (44) aan een eindige bovengrens gebonden blijft, moet elke bijdrage aan (44) beperkt blijven en dus ook de bijdrage die ontstaat door alle termen die ω', ω'', ω''', ..., bevatten, te verwaarlozen. Door de bedoelde bijdrage af te zonderen kan men voor (44) schrijven,

$$|j_g| \leqq a_1 \sum_{n=0}^{\infty} |\omega|^n (a_3 + a_2 |\frac{d}{d\delta}|)^{(n)} \|E\| + \text{een voor } \omega' = 0 \tag{47}$$

verdwijnende functie.

Opgemerkt wordt dat het gedrag van ω', ω'', ω''', ..., sterk af zal hangen van de grootte van het polaire traagheidsmoment j_p en het is niet onredelijk te onderstellen dat de eerste term van (47) voor grote j_p de belangrijkste bijdrage zal vormen voor de maximale excursies van ω. Nu geldt

$$\sum_{n=0}^{\infty} |\omega|^n (a_3 + a_2 |\frac{d}{d\delta}|)^{(n)} \|E\| = \sum_{n=0}^{\infty} |a_3 \omega|^n \left[|\frac{d}{d\delta}|^{(0)} + \right.$$

$$\left. + \binom{n}{1}\left(\frac{a_2}{a_3}\right)|\frac{d}{d\delta}|^{(1)} + \cdots + \binom{n}{n}\left(\frac{a_2}{a_3}\right)|\frac{d}{d\delta}|^{(n)} \right] \|E\| =$$

$$= \sum_{n=0}^{\infty} |a_3 \omega|^n \left[(1 + \frac{a_2}{a_3})^{(n)} + (1 + e_f^{c2})^{\frac{1}{2}} - 1 \right] \tag{48}$$

een bijdrage die slechts eindig blijft voor

$$|\dot{\delta}| = |\omega| < \frac{1}{a_2 + a_3} \approx \frac{\alpha_f}{3 + 2^{\frac{1}{2}} \alpha_f} \tag{49}$$

Voor $\alpha_f = 0{,}15$ (zie tabel 1) zal dus (34) alleen voor $|\dot{\delta}| < 0{,}05$ kunnen convergeren.

De moeite die men aan analytische arbeid moet besteden om benaderde vergelijkingen voor het stelsel (1) te verkrijgen is, ten opzichte van het gemak waarmede men met behulp van een analogonrekentuig een goed inzicht in het gedrag van het niet benaderde stelsel verkrijgt, onevenredig groot en geeft bovendien resultaten waarvan de nauwkeurigheid nauwelijks valt te schatten.

Literatuur

[1] L. Dreyfus: Elektrotechnik und Maschinenbau, 1911, 323 e.v.
[2] Th. Bödefeld: E. und M. 1930, 689 e.v.
[3] H. E. Edgerton, F. J. Zak: Journ. I.E.E. 1930, 1205 e.v.
[4] H. E. Edgerton, P. Fourmarier: A.I.E.E. 1931, 769 e.v.
[5] W. V. Lyon, H. E. Edgerton: A.I.E.E. 1930, 686 e.v.
[6] A. Timascheff: Stabilität elektrischer Drehstrom-Kraftübertragungen, Diss., 1940, Berlin.
[7] D. R. Shoults, S. B. Crary, A. H. Lauder: A.I.E.E. 1935, 1385 e.v.
[8] Th. Laible: Die Theorie der Synchronmaschine im nichtstationären Betrieb. 1952, Berlin.
[9] Concordia: Synchronous Machines, Theory and Performance. 1951, New York.
[10] N. Vlasov: zie bespreking in het boekwerk van [11].
[11] N. Minorsky: Introduction to Non-Linear Mechanics. J. W. Edwards 1947, Ann Arbor, 124-130.
[12] J. J. Stoker: Nonlinear Vibrations. Interscience Publ., 1950, New York, 67-80.
[13] W. A. Coppel: On the Equation of a Synchronous Motor. Quart. J. of Mech. and Appl. Math., 12, part 2 (May 1959), 242-256.
[14] Vakantieleergang Elektrotechniek 1959. Overdruk uit het Weekblad *De Ingenieur*, 1960, nr. 25, 31, 33, 38, 41 en 43.
[15] D. Köppel: Berechnung des zeitlichen Verlaufs der Stabströme im Dämpferkäfig einer Synchronmaschine mit Vollpolläufer für beliebige Betriebszustände. Diss., Karlsruhe (TH), 1968.
[16] R. Bellman: Stability Theory of Differential Equations, McGraw-Hill, 1953, 1-9.

530.12

A note on unified field theory



bij prof. dr. R. Timman, Delft University of Technology, Department of General Sciences, Subdepartment of Mathematics

Summary: A unified field theory is considered, where the electromagnetic field is represented by the anti-symmetric part of a non-symmetric connexion.

Most attempts at a unified field theory (e.g. V. Hlavatý [1]) introduce a non-symmetric metric tensor and relate the gravitation to the symmetrical part and the electromagnetic properties to the antisymmetrical part. The first gives an ordinary Riemann space and its tensor is usually used to form covariant vectors. Apparently both tensors are, right from the beginning, completely separated. The existence of electrodynamics in special relativity, where the space-time continuum has a constant metric tensor, while the electromagnetic field is governed by the antisymmetric field tensor leads to a concept, where the metric tensor is symmetrical, but the connexion Γ^i_{jk} has an antisymmetric part

$$S^i_{jk} = \tfrac{1}{2}(\Gamma^i_{jk} - \Gamma^i_{jk})$$

By inspection of the transformation laws for the connexion (e.g. Schouten [2]) this object appears to be a tensor, which is called the *torsion* of the space.

Here we shall derive a Hamilton principle in this space, which allows us to postulate a relation between the torsion

tensor and the electromagnetic tensor. A similar consideration in completely different notation is given by Sciama [3].

The parallel displacement of a vector v^i, which is transferred from a point x^i to a point $x^i + dx^i$ in the space with holonomous coordinates x^i is defined as a displacement for which

$$\delta v^i = dv^i + \Gamma^i_{jk} v^k dx^j = 0$$

The object Γ^i_{jk} is called the connexion in the space and the *vector* δv^i is called the covariant differential of v^i. It is a linear homogeneous function of the displacement dx^i and it satisfies the ordinary rules with respect to addition and multiplication of a differential. As was mentioned above Γ^i_{jk} need not be symmetrical and

$$S^i_{jk} = \tfrac{1}{2}(\Gamma^i_{jk} - \Gamma^i_{kj})$$

is a tensor, the significance of which we shall investigate. If in a point x^i we have two line-elements dx^i and δx^i and we displace δx^i along the line-element dx^i, we have for a parallel displacement the endpoint

$$x^i + dx^i + \delta x^i + \Gamma^i_{jk} \delta x^k dx^j = x^i + dx^i + \delta x^i + d\delta x^i$$

on the other hand for a parallel displacement of the vector dx^i along the vector δx^j we have

$$x^i + \delta x^i + dx^i + \Gamma^i{}_{kj} \delta x^k dx^j = x^i + \delta x^i + dx^i + \delta dx^i$$

In our case the endpoints do not coincide, but they differ by a vector

$$d\delta x^i - \delta dx^i = (\Gamma^i_{jk} - \Gamma^i_{kj}) \delta x^k dx^j = 2 S^i_{jk} \delta x^k dx^j.$$

This property will be used for the derivation of a Hamilton principle.

$$\int_A^B ds = \text{stationary}$$

for different curves, running from a fixed point A in space-time to a fixed point B. Here ds is the line-element, defined by the symmetric metric tensor g_{ik}

$$ds^2 = g_{ik} dx^i dx^k$$

It can be shown (Schouten [p. 132, 2]), that, if we put the covariant derivatives

$$\nabla_i g_{jk} = 0$$

that the tensors g_{ik} and S^i_{jk} define uniquely the connexion.

$$\Gamma^i_{jk} = \{^{\,i}_{jk}\} + S_{jk}^{\cdot\cdot i} - S_{j\cdot k}^{\cdot i} + S^i_{\cdot jk}$$

where the tensors $S_{j\cdot k}^{\cdot i}$ and $S^i_{\cdot jk}$ are derived from $S_{jk}^{\cdot\cdot i}$ by raising and lowering of indices

$$S_{j\cdot k}^{\cdot i} = g^{il} g_{km} S_{jl}^{\cdot\cdot m};\ S^i_{\cdot jk} = g^{il} g_{km} S_{lj}^{\cdot\cdot m}$$

and $\{^{\,i}_{jk}\}$ stands for the Christoffel symbol

$$\{^{\,i}_{jk}\} = \tfrac{1}{2} g^{il} \{\partial_j g_{kl} + \partial_k g_{jl} - \partial_l g_{jk}\}$$

Consider now in this space the variation of an integral

$$\int_A^B L(x^i, u^j)\, ds$$

where $u^i = \dfrac{dx^i}{ds}$. In each point $x^i(s)$ of the extremal we define a variation $\delta x^i(s)$ with a corresponding $\delta u^i = \delta \dfrac{dx^i}{ds}$

From the relation derived above, we can infer that, remarking, that $\dfrac{d}{ds}$ means a differentiation along the line-element dx^i

$$\delta u^i = \delta \frac{dx^i}{ds} = \frac{d}{ds} \delta x^i - 2 S^i_{jk} \delta x^k \frac{dx^j}{ds}$$

The condition for stationarity of the integral now reads

$$\int_A^B \delta L\, ds = \int_A^B \left\{ L(x^i + \delta x^i, u^i + \delta u^i)\, L(x^i, u^i) \right\} ds =$$

$$= \int_A^B \left(\frac{\partial L}{\partial x^i} \delta x^i + \frac{\partial L}{\partial u^i} \delta u^i \right) ds \geqq 0$$

for arbitrary variations δx^i which vanish at the endpoints.

Substitution of the expression for δu^i gives

$$\int_A^B \left(\frac{\partial L}{\partial x^i} \partial x^i + \frac{\partial L}{\partial u^i} \frac{d}{ds} x^i - 2 \frac{\partial L}{\partial u^i} S^i_{jk} \partial x^k u^j \right) ds \geqq 0$$

or after partial integration of the second term,

$$\int_A^B \left(\frac{\partial L}{\partial x^i} - \frac{d}{ds} \frac{\partial L}{\partial u^i} - 2 S^j_{kl} \frac{\partial L}{\partial u^j} u^k \right) x^i \geqq 0$$

or

$$\frac{d}{ds} \frac{\partial L}{\partial u^i} - \frac{\partial L}{\partial x^i} + 2 \frac{\partial L}{\partial u^j} S^j_{ki} u^k = 0$$

We now remark, that $L = \sqrt{g_{ik} u^i u^k}$, which, along the extremal has the value 1 and carry out the differentiation. The first two terms give, similar results as those for a symmetrical connection, but the last term gives an addition.

$$g_{il} \frac{du^l}{ds} + \left\{ jk, l \right\} u^j u^k - 2 S^k_{jl} g_{kl} u^l u^j = 0$$

From this result we find the equation for the geodetic line

$$\frac{du^i}{ds} + \Gamma^i_{jk} u^j u^k = 0$$

For the equation of motion of an electrical particle in a gravitational field we have (Landau and Lifschitz [4])

$$\frac{du^i}{ds} + \left\{ ^{\,i}_{jk} \right\} u^j u^k = \frac{1}{mc^2} F_{jk} g^{ji} u^k$$

where F_{jk} is the electromagnetic field tensor.

This leads to a tentative identification

$$\frac{1}{mc^2} F_{jk} = S^i_{jk} u_i$$

Regarding the formula for the vector potential A_k we have

$$F_{jk} = \nabla_j A_k - \nabla_k A_j = \partial_j A_k - \partial_k A_i - \Gamma^i_{jk} A_j + \Gamma^i_{kj} =$$

$$= \partial_j A_k - \partial_k A_j - 2 S^i_{jk} A_i$$

This shows, that in the case, of a flow field, which satisfies the relation

$$\partial_j u_i - \partial_i u_j = 0$$

we may indentify u_i with the vector potential

$$A_i = \frac{mc^2}{l} u_i$$

It might be of some interest to work out further consequences of this hypothesis.

[1] V. HLAVATY: Geometry of Einstein's unified field theory. Groningen, 1957.
[2] J. A. SCHOUTEN: Ricci Calculus. Berlin, 1954.
[3] D. W. SCIAMA: Recent Developments in general relativity. Warsawa, 1962.
[4] L. LANDAU et E. LIFSCHITZ: Théorie du champ. Moscou, 1962.

Uit het NERG

Administratie van het NERG: Postbus 6108, Den Haag.
Giro 94746 t.n.v. penningmeester NERG, Den Haag.
Secretariaat van de Examencommissie-NERG: van Geusaustraat 151, Voorburg.

Ledenmutaties

Voorgestelde leden
Ir. C. A. Bruijnis, van Adrichemstraat 31, Delft.
Dr. ir. J. B. H. Peek, Dennenlaan 10, Nuenen.
Ir. R. van der Pol, Tongelresestraat 356, Eindhoven.

Nieuwe leden
Ir. H. G. H. Gooren, Maasstraat 43, Arcen.

Nieuwe adressen van leden
Ir. F. H. Altmann, Hatishbi-street 25, Haifa, Israel.
H. M. J. Bucx, Stephensonstraat 72, Den Haag.
Ir. W. F. Dil, Argonautenlaan 24, Eindhoven.
R. Y. Drost, Ing., Diepenbroickpark 26, Weesp.
Ir. J. H. Hendriks, Symon Blaauwboerstraat 40, Den Helder.
Ir. L. L. Kossakowski, Pettelaarseweg 549, Den Bosch.
Ir. A. S. T. Kruijf, Dr. H. Colijnlaan 170, Rijswijk Z.H.

Overleden
Ir. A. H. de Voogt, Piet Heinstraat 62a, Den Haag.